

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:35:21 ; Search time 42 Seconds  
(without alignments)  
1278.308 Million cell updates/sec

Title: US-09-189-415b-11  
Perfect score: 2840  
Sequence: 1 MPIGNLGHNPVNNVNSIPPAP.....SNSAVNTSNPPAGSHRFV 558

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2840	100.0	558	2	A98199	translocated intim
2	2840	100.0	558	2	E86045	probable transloca
3	200	7.0	2232	2	T34434	hypothetical prote
4	195.5	6.9	1229	2	T25697	hypothetical prote
5	187	6.6	1291	2	T13389	hypothetical prote
6	172.5	6.1	1630	2	A53577	ascites sialoglyco
7	172.5	6.1	2271	2	F90073	hypothetical prote
8	167.5	5.9	1192	2	T18611	probable serine/th
9	167	5.9	2468	2	A83412	hypothetical prote
10	166	5.8	796	2	T21460	hypothetical prote
11	164	5.8	971	2	B90835	hypothetical prote
12	164	5.8	973	2	C85693	probable tail fibe
13	164	5.8	1275	2	T33369	probable membrane
14	164	5.8	1770	2	A71517	hypothetical prote
15	163.5	5.8	1246	2	G89287	protein H39E23.1
16	163.5	5.8	1802	2	S69703	HKR1 protein precu
17	163	5.7	786	2	T16509	hypothetical prote
18	160	5.6	1122	2	G64887	probable tail fibe
19	160	5.6	1829	2	T24583	hypothetical prote
20	158.5	5.6	1063	2	D86731	hypothetical prote
21	157.5	5.5	3013	2	A80480	probable invasin y
22	156.5	5.5	3570	2	T34525	mucin MUC5B, trach
23	155.5	5.5	3507	2	T34513	hypothetical prote
24	152.5	5.4	1077	2	A44067	serine-rich protei
25	152	5.4	461	2	JN0097	secreted 45K prote
26	152	5.4	997	2	T43523	cut17 protein - fi
27	151.5	5.3	2660	2	E85822	probable invasin z
28	151	5.3	918	2	T02759	hypothetical prote
29	151	5.3	1306	2	S25370	MSB2 protein - yea

30	150	5.3	888	2	T46726	secreted acid phos
31	150	5.3	1026	1	A40315	maternal effect pr
32	150	5.3	1034	2	JC2143	ice nucleation act
33	150	5.3	1460	2	D81675	polymorphic membra
34	149.5	5.3	1772	2	A45532	major merozoite su
35	149.5	5.3	2514	2	F81045	hemagglutinin/hemo
36	148.5	5.2	770	2	T51024	related to C2H2 zi
37	148.5	5.2	1189	2	S56852	hypothetical prote
38	148	5.2	2035	2	A40718	host cell factor C
39	148	5.2	3190	2	T13828	CREB-binding prote
40	147.5	5.2	797	1	VBEX1	glycoprotein X pre
41	147.5	5.2	1547	2	T28657	blackjack protein,
42	147.5	5.2	2055	2	T31110	extracellular matr
43	147.5	5.2	2059	2	D82671	surface protein XF
44	147.5	5.2	3624	2	AD0835	large repetitive p
45	147.5	5.2	4558	2	C82199	RTX toxin RtxA VCI

ALIGNMENTS

RESULT 1

A:98199  
translocated intimin receptor Tir [imported] - Escherichia coli (strain O157:H7, substrai  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Jul-2004  
C:Accession: A98199  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: A98199  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-558 <HAY>  
A:Cross-references: UNIPROT:Q9R396; GB:BA000007; PIDN:BA037984.1; PID:gl3364036; GSPDB:G  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: EC94561

Query Match	100.0%	Score 2840;	DB 2;	Length 558;
Best Local Similarity	100.0%	Pred. No. 3.4e-149;		
Matches 558;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPIGNLGHNPVNNVNSIPPAPLPSTQDGAGRGQLINSTGPIGSRALFTPVNSMADSGD	60	
Db	1	MPIGNLGHNPVNNVNSIPPAPLPSTQDGAGRGQLINSTGPIGSRALFTPVNSMADSGD	60	
Qy	61	NRASDVPGLPVNPMLAAASEITLNDGFVLDHGHPLDNLNROIIGSSVFRVETQEDGKHIA	120	
Db	61	NRASDVPGLPVNPMLAAASEITLNDGFVLDHGHPLDNLNROIIGSSVFRVETQEDGKHIA	120	
Qy	121	VGQRNGVETSVVLSDQEVARIQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQITL	180	
Db	121	VGQRNGVETSVVLSDQEVARIQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQIL	180	
Qy	181	ELLEPKGTGSGKAGESKGVGELRESNSGAENTTETQSTSSLSRSDPKLWALGTVAT	240	
Db	181	ELLEPKGTGSGKAGESKGVGELRESNSGAENTTETQSTSSLSRSDPKLWALGTVAT	240	
Qy	241	GLIGLAATGIQVALALTPEDSPPTTDDPAAASATETATRDLTKEAFQNPQNKVNIDE	300	
Db	241	GLIGLAATGIQVALALTPEDSPPTTDDPAAASATETATRDLTKEAFQNPQNKVNIDE	300	
Qy	301	LGNAPSVGLKDDVVANIEEQAAGEAKQQAENNAQAQKDYEQQAQKEELKVS	360	
Db	301	LGNAPSVGLKDDVVANIEEQAAGEAKQQAENNAQAQKDYEQQAQKEELKVS	360	
Qy	361	AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTTSARTVENKPNNTPAQ	420	
Db	361	AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTTSARTVENKPNNTPAQ	420	



hypothetical protein F16F9.2 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Acession: T25697  
R:Fulton, B.  
submitted to the EMBL Data Library, August 1996  
A/Description: The sequence of C. elegans cosmid F16F9.  
A/Reference number: Z20071  
A/Acession: T25697  
A/Status: preliminary; translated from GB/EMBL/DDBBJ  
A/Molecule type: DNA  
A/Residues: 1-1229 <UL>  
A/Cross-references: UNIPROT:Q94185; EMBL:U67956; PIDN:AAB07691.1; GSPDB:GN00028; CESP:F16F9  
A/Experimental source: strain Bristol N2; clone F16F9  
C/Genetics:  
A/Gene: CESP:F16F9.2  
A/Map position: X  
A/Introns: 35/1; 361/1; 384/1; 482/2; 517/1; 971/1; 1021/2; 1179/3

Query Match            6.9%   Score 195.5; DB 2; Length 1229;  
Best Local Similarity 22.0%; Pred No. 0.0048;  
Matches 90; Conservative 54; Mismatches 155; Indels 111; Gaps 13;

Qy 167 TVASDITEARQRILELLEPRTGKSGAGSKGVGLR-----ESNSGAENETIOTSTST 222  
     :::: : :: : : : : : : : : : : : : : : : :  
Db 597 SVSTEKTSTTKASTHEEPTTDEPTTHSSITGATTPELSTTSEETTTTELKITTEG 656  
     :  
Qy 223 SSIRSDPKMLALGTWATGLIGLAATGIQVALALTPEPPSPPTTDPDAASAATETARQ 282  
     :  
Db 657 STTTEEP-----TTTAIFAFAESTGII---TTDEETTSTTTPETSTKEIVTESA 704  
     :  
Qy 283 LTKEAF-----QNPDKNQINDELGNAIPSGVLKDDV-VANIEEQAKAGEAAQQ 332  
     :  
Db 705 ITQTSVSWBESSPTOLPERKWAIVNKFNHLE--VLKEKKRLIKEKESTSTTGSDS-- 760  
     :  
Qy 333 AIENNAQAQKYDPAQKROELKVSSGAGYGLSGALLGGYGIVATAALHRKNQPVEQ 392  
     :  
Db 761 --FTTVVAENIDVTYTEKEV-----VOTPIITEKSTTQEE 797  
     :  
Qy 393 TTTTTTTTTTSARTVENKANPNTPAQGNVDTPGS EDTMESRRSSMASTSPFDTSSTG 452  
     ||||| :  
Db 798 TTTTTTTEKTSKTTTEKETTSESA-----TTTETSSEPTTESTTVDTSAT 846  
     :  
Qy 453 TVQNPYA--DVKTSLHDSQVPTSNS-----NTSVQNMG----- 483  
     :  
Db 847 TEESSTAETTTTSAETSETTSESAAFITESAPENTALQS SSKSEENESSAEKPGARR 906  
     :  
Qy 484 -----NTDSVVYSTIOHPDRDTTDNGARLLGNPSAGIQST 518  
     :  
Db 907 DFPVKKHKTVPKAETTSAVAASSTTETPTTTTEKSTLTLETTP---IEAT 953  
     :

RESULT 5  
T13389 hypothetical protein 115C2.10 - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: T13389  
R:Salles, C.; Valenti, P.; Darlamitsou, A.; Henderson, N.; Campbell, L.; Glover, D.  
submitted to the EMBL Data Library, May 1999  
A:Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.  
A:Reference number: Z17665  
A:Accession: T13389  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1291 <CAT>  
A:Cross-references: UNIPROT:O77261; EMBL:AL031581; NID:e1320978; PID:e1426292; PIDN:CAA2  
C:Genetics:  
A:Cross-references: FlyBase:FBgn0020381  
A:Map position: X  
A:Introns: 238/3; 1225/1  
A>Note: EG:115C2.10

Query Match	6.6%; Score 187; DB 2; Length 1291;
Best Local Similarity	20.4%; Pred. No. 0.015;
Matches	130; Conservative 74; Mismatches 245; Indels 188; Gaps 25;
Qy	5 NLGHNPVNNVSIIPAPPLPSQTGG---AGRGQLINSTGPLGRALFTPVNNSMADSGDN 61
Db	438 NTSNGNSNTNDSTGPGSETSTNGLVASGGAG-----GATGAMLPTP---SQOSTGGK 488
Qy	62 RAS-----DVFGLPVNPMRLA-----ASEITLNDGFEVLHDHGPLDLTLNRQ 102
Db	489 EATAAVALLEKKLPVNVVSLPTKMLRQKMTKYDAEMINAAAYOOHHH-----539
Qy	103 IGSVFRVETQEDGKHIAVGORNGVFSTVVLSDQEARLQSIDPEGKDKFVFTGGRGAG 162
Db	540 --QHFFHHHHHHHHHGHGHAAGTGAETAAVQOQMAAQKPG-----VGGTGAAG 588
Qy	163 HAMVTVASDITEARQRIELLEPKGTGESKAGESKVGELRSNSGAENTTQTSTST 222
Db	589 NAGATTVSSVA-----AGAGSEVNGRSTSLRKSMTVRNS-----TSSSI 622
Qy	223 SSLSRDPKMLALGTVATGLIGLAATGIVQALATPEPDSPTTTDPAASAATATARDQ 282
Db	628 STASADEVI-----APVVAASISLPSKAPVLMRCKPAQMAIALHQ 670
Qy	283 LTKEAFQNPQNKVINDLGNALPSGYLKDQDVANIEEQAAAGEEAKQ-----A 333
Db	671 SQOQLRRSRQEKLTJ-----GESSD---TSSEQKK-----EQQDHDHLPQKWF 716
Qy	334 IENNAQAKYDQQAQRELBKVVSGAGYGLSGALILGGIGIVAVTAALHRKKNQPVQ 393
Db	717 LABEPOPEKS-EBKQEQQRKVRNRSAGRYGL-----VARLATAHNN-----IA 760
Qy	394 TTTTITTTTTTSARTV---ENKPANNTPAQGNVDTPGSEDTEMSRRSSMASTSTFPDTSS 450
Db	761 TTTNSSSSSNKATITTCNNHNSNNSRINHNSLSRLSVKSRKPAPSEASSIPSSSTSS 820
Qy	451 IGTVQ-----NPYADVKTSLHDSQVPTSNSTSVQNMGNMTDSVVYSTIQH-----495
Db	821 ENQQQATRRSCSPPTPAYKKNLLASFPDPFPSTQGIKEQLKDESVTYSVVKQKSRRAA 880
Qy	496 -----PPRTDTNGARL-LGNPSAGTST-----YARLALSGG 527
Db	881 LAAQSIHCALGGFPTGTSQSKRAQAAGEPTTSCSSTTISNVEPLKTPERRKLTLR 940
Qy	528 LRH-----DMGGLTGGSNAVNTSNPPAPGSHR 556
Db	941 MKRSPILDEVIELG--TSLNNGGARG-----APGSHR 971
RESULT 6	
A53577	ascites sialoglycoprotein 1 - rat (fragments)
C;Species:	Rattus norvegicus (Norway rat)
C;Date:	12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 07-Feb-1997
C;Accession:	A53577
R;Wu, K.; Fregien, N.; Carraway, K.L.	
J. Biol. Chem.	269, 11950-11955, 1994
A;Title:	Molecular cloning and sequencing of the mucin subunit of a heterodimer
A;Reference number:	A53577; PMID:94216302; PMID:8163496
A;Accession:	A53577
A;Status:	preliminary
A;Molecule type:	mRNA
A;Residues:	1-1630 -WUA-
A;Cross-references:	GB:U06746
C;Keywords:	glycoprotein
Query Match	6.1%; Score 172.5; DB 2; Length 1630;
Best Local Similarity	20.0%; Pred. No. 0.13;
Matches	136; Conservative 70; Mismatches 272; Indels 203; Gaps 23;
Qy	12 VNSIIPAP-----PLPSQTDGAGRGQLINSTGPLGRALFTPVNNSMADSGDNRA 67
Db	28 VNTSTTSAPKTSIALPSSTNPQSVNPTA-----SSRYMTKNTGOASPMTVSSIT 82



A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA1874

Query Match	5.9%;	Score 167;	DB 2;	Length 2468;
Best Local Similarity	24.3%;	Pred. No. 0.46;		
Matches 151;	Conservative	53;	Mismatches 250;	Indels 162; Gaps 34;

  

QY	2	PIGNLGHNNVNNISIPAPPLPQSDTGAGRGQLINSTGLSGRALFTFVRNSMADSGDN	61
DB	701	PIQGVATDGSNGWSFTTPLPN-----GTVNATATDAS-----GNTSAGSSVT	745
QY	62	RASDVPLPV-NPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVRFVETQEDGKHIA	120
DB	746	VDSVAPATPVINF-----SNGTILSGTAEP-----GSSV-----TLTDGNGNP	783
QY	121	VGORN-----GVETSVULSDOEYARLOSIDPEGKDKFVTCGRGGA---GHAMVTVAS	170
DB	784	IGQVTADGSNGWSFTTSTPLADGTVVNATATDAGN-----TSQGSSTTVDGVAPTTFTV	838
QY	171	DITEARQRIELELLEPKGTGESKGA-----GESKGVGELRESNGAENTTETQTSTSTSSL	225
DB	839	NLSNGSSL-----SGTAEPGSTVILTDGNGNPIAEVTDGSG--NWTVT-----	880
QY	226	RSDPKLWLALGTATGLIGLAATGIVQALALTEPPDSP-----TTTDPDAASAT	275
DB	881	---PSTPIANGTVVNVVQAADAGNSSPGASVTVDSQAPAPVVPNSNGTTLTGTAEPGAT	937
QY	276	ETATRDQLTKFAPQNPDPNQKVNIDELGN--AIPSGVLKDDVANIIEQAKAAGEEAKQQA	333
DB	938	VTLT-----DGNNGNPIGQ-VTADGSGNWSFTTGTPLANGTVVN-----ATASDPT	981
QY	334	IENNAQAKKYDQQAQKQBEELKVSSGA---GYGLSGLALI-LGGGICGAVTAALHRKNQP	389
DB	982	GNTSAPASTTVD--SVAPAAPVVPNSNGAEISGTAEPCGATVTLTDGSG-----NP	1029
QY	390	VEGTTT-----TTTTTTTTTSARTVEN-----KPNANTPAQGN--VD--TPGESDTEMSRR	435
DB	1030	IGQVTADGSNGWSFTTSTPLADGTVVNATATDAGTGTGGGSTTVDAIAPATTVNLNSG	1089
QY	436	SSMASTS---STPFDTSSIGTVQNPYADVKTS-----LHDSQVPTSNS---NTSVQMGNN	484
DB	1090	SSLSGTAEPGSTVILTDGNG--NPIAEVTDGSGNWTVPSTPIANGTVVNVVQAQASG	1146
QY	485	TDS-----VYYSITQHPPRDTTNGARLLGNPSAGIQSIVARIALSGELRHDGGLTG-G	538
DB	1147	NSSPPATVTVDDSAPPAPVNPNSGVVISGTAEAGATVT---LTDAGG--NPIGQVTDAGD	1201
QY	539	SNSAVNTSNPPAPGS	554
DB	1202	SGNWSFTTGTPLANGT	1217

RESULT 10

hypnotical protein ZK945.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T21460  
R:Wilkinson, J.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: Z19425  
A:Accession: T21460  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-796 <WIL>  
A:Cross-references: EMBL:Z48582; PIDN:CAA88469.1; GSPDB:GN00020; CESP:ZK945.10  
A:Experimental source: clone F27ES  
C:Genetics:  
A:Gene: CESP:ZK945.10  
A:Map position: 2  
A:Introns: 62/3; 124/3; 169/1; 216/1; 262/3; 371/1; 581/3; 608/1

Qy	182	LLEPKGTGSKGAGESKGVGLRSLNSNGAENTTTTQTSTSTLSLRSDPKLWLALGTVATG	241
:::		:::	
Db	155	-----ASTSAGQAASSAQSSASCTASTTKATEASKSAAAEESK-----SAAAT	199
:::		:::	
Qy	242	LIGLAATGIQVALALTPEPDSPTITDPPDAASAETATETROLTLKEAFONPDNOKVIMDEL	301
:::		:::	
Db	200	SAGAAKTSETNAAVSQSOAATSASTATTKASEAASSARDASAKSEAAKSSETSAA--	255
:::		:::	
Qy	302	GNAIPSGVLKDWDVANIEEQAAAAGEBAK-QOAIENNAQAQKKYDEQQAQRQBELKVSSG	360
:::		:::	
Db	256	-----SASSAASSATAGNNSKAAKTSETNAKSSETAEQSASAAAAAGSK--TA	301
:::		:::	
Qy	361	AGYGILGALIIGGGIVGVAVTAAHLHRKNQPVEQTITTTTTTSARTVENKPANNTPAQ-	419
:::		:::	
Db	302	AALSASAASTSAGQAASATAA---GKSABESAASSASTATTKAGATEAQASAAASASA	357
:::		:::	
Qy	420	--GNVDTPGESDTEWRRSRMSASTSTFTFDDTISGITVQNPHYADVKLTSHDSQVPITSNN	476
:::		:::	
Db	358	AKTSETNAKGSTSAESSKTAAASASASASASSASASASAKDEATRQASAAKSATTASTK	417
:::		:::	
Qy	477	TSVQNMGNITDVSVVYSITQHPRDRITTDNGA	505
:::		:::	
Db	418	A-----TEAGSATAAQAQSKSTAESAA	439
RESULT 12			
C85693	probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli		
C:Species:	Escherichia coli		
C>Date:	16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004		
C:Accession:	C85693		
R;Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.			
iller, L.; Grothebeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamousis, K.			
Nature 409, 529-533, 2001			
A:title:	Genome sequence of enterohemorrhagic Escherichia coli O157:H7.		
A:Reference number:	A85480; MUID:21074935; PMID:11206551		
A:Accession:	C85693		
A>Status:	preliminary		
A:Molecule type:	DNA		
A:Residues:	1-973 <ST>		
A:Cross-references:	UNIPROT:P8XDQ4; GB:A5005174; NID:g12514847; PIDN:AAG56007..		
A:Experimental source:	strain O157:H7, substrain EDL933		
C:Genetics:			
A:Gene:	Z1918		
Query Match            5.8%; Score 164; DB 2; Length 973;			
Best Local Similarity     21.2%; Pred. No. 0.19;			
Matches      95; Conservative          66; Mismatches       198; Indels       90; Gaps       15			
Qy	79	SSIINDGFVLH-----DHGPDLTNRQTGSVFVRVETOEDKHIAVGQ-----	123
:::		:::	
Db	61	SVILLVEGFPSPSHACTITVVEDSQP-GTLNDFLGAM-----TEDDVREPALRRFFELMVEE	114
:::		:::	
Qy	124	--RVGVETSVMVLDQEVARLOSIDPEGKDFVTGGRGGGHAWTVVASDITEARORILE	181
:::		:::	
Db	115	VARN---ASAQAQNTAAAKGASDAQS-----TSAREAAATHA-TDAADSARA-----	156
:::		:::	
Qy	182	LLEPKGTGSKGAGESKGVGLRSLNSNGAENTTTTQTSTSTLSLRSDPKLWLALGTVATG	241
:::		:::	
Db	157	-----ASTSAGQAASSAQSSASCTASTTKATEASKSAAAEESK-----SAAAT	201
:::		:::	
Qy	242	LIGLAATGIQVALALTPEPDSPTITDPPDAASAETATETROLTLKEAFONPDNOKVIMDEL	301
:::		:::	
Db	202	SAGAAKTSETNAAVSQSOAATSASTATTKASEAASSARDASAKSEAAKSSETSAA--	257
:::		:::	
Qy	302	GNAIPSGVLKDWDVANIEEQAAAAGEBAK-QOAIENNAQAQKKYDEQQAQRQBELKVSSG	360
:::		:::	
Db	258	-----SASSAASSATAGNNSKAAKTSETNAKSSETAEQSASAAAAAGSK--TA	303
:::		:::	
Qy	361	AGYGILGALIIGGGIVGVAVTAAHLHRKNQPVEQTITTTTTTSARTVENKPANNTPAQ-	419
:::		:::	
Db	304	AALSASAASTSAGQAASATAA---GKSABESAASSASTATTKAGATEAQASAAASASA	359

Query Match 5.8%; Score 166; DB 2; Length 796;  
Best Local Similarity 20.3%; Pred. No. 0.11;  
Matches 108; Conservative 89; Mismatches 223; Indels 112; Gaps 19;

QY 68 GLPNPMLAASETLND-----GFEVLHDHGPLDLNRQIGSSVFRVETQEDGKHI 119  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
137 GLFNSWTWITLNEVDDDEISIAVEAKYECYDDG----IDRCDSLWW---IQVGNGEM 189  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
120 A-VGRNGRVTSVVLSDOEVARLOSIDPEKGDKVFFTGCRGGAGHAMVTVASDITEARQR 178  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
190 ALLGVREKESGEI--NEEYARRMCKRPYRSEK-----STAISDSQGV 230  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
179 ILELLPKGTGESKAGSKSGVKGLRSNKAENNTTQTSTSTSLRSRDPKWLALGTV 238  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
231 YYDGCVLKGVRAKQFSVRTSGSPTLRMRKRDAKDNTCDYTIESTSTSTTTTPTTTVTSTV 290  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
239 -ATGLIGLAATGIVOALALTPEPPSPITT-----DPDAASATATATEDQLTKEAFQ 289  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
291 TSTTTVTSTSVTAMST 350  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
290 NPDNQKNVIDELGNAIPEGVLKD--DVVANIEOAKAAG-EAAKQAIIENNAOAQKYDE 346  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
351 SPSS-----TTLSTSIPTTTTPEITSLLSSLPDNAICSYLDEITTSFTFTTMLTSTTTE 405  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
347 QQAQKBELKVSSGAGVLSGALLGGIGAVVAALHRKNQPVEQTTTTTTTTTTTT--S 404  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
406 EPSTSTTTTTEVTS-----TSTVTPTTTLTSTASTSTSTPS 445  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
405 ARTVENKPA-----NNTPAQGVDTPGSEDPTMESRRSSMASTSTSTFDFTSSIGTVQ 455  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
446 TSTVTTSPSTPVSTVTSSSSSSTVTTTFSTESTSTSPSSVTTSTTAPSTSTTGPS 505  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
456 NPYADVKTSLHDSQVPSTNSGNTSDSVVYSTIOHPRTDTPONGARLLGNP---- 511  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
506 S-----SSTPSSTASSVSSTASSTQSSTTQOSTTKSETTSSDGT---NPDFYF 555  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
512 -----SAGIQSTYARLALSGLRHDMGLGGSNSAVNTSNPPAPQSHRFV 558  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
556 VEKATTTYDYDSTSVNLNLSGL-----GIIGQYTSIECTS-----PTSSNYIV 597

RESULT 11  
B90835  
Probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain RIM C); Species: Escherichia coli  
C; Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C; Accession: B90835  
R; Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasawara, T.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
A; Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference: UNIPROT\_Q8XDK4; GB:BA000007; PID:NBAR35073.1; PID:g13361114; GSPDB:G A; Reference number: A99629; MUID:21156231; PMID:11258796  
A; Accession: B90835  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-971 <HAY>  
A; Cross-references: UNIPROT\_Q8XDK4; GB:BA000007; PID:NBAR35073.1; PID:g13361114; GSPDB:G A; Experimental source: strain O157:H7, substrain RIMD 0509952  
C; Genetics:  
A; Gene: ECa1650

Query Match 5.8%; Score 164; DB 2; Length 971;  
Best Local Similarity 21.2%; Pred. No. 0.19;  
Matches 95; Conservative 66; Mismatches 198; Indels 90; Gaps 15;

QY 79 SETLNDGFVELH-----DHGFLDLNRQIGSSVFRVETQEDGHIAVGQ----- 123  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
59 SVILLVEGFPFHAGTITVVEDSQP-GTLNDFLGAM-----TEDDVREALRRFELMVEE 112  
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
124 --RNVETSVVLSDOEVARLOSIDPEKGDKVFFTGCRGGAGHAMVTVASDITEARORILE 181  
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
113 VARN---ASAVALQNATAAKKSADAS-----TSAREAAATHA--TDAADSARA----- 154





A>Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: G89287  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1246 <STO>  
A:Cross-references: GB:chr\_V; PIDN:CAB09532.1; PID:G3878100; GSPDB:GN00023; CESP:H39E23.  
C:Genetics:  
A:Gene: H39E23.1  
A:Map position: 5

```
Query Match          5.8%; Score 163.5; DB 2; Length 1246;
Best Local Similarity 19.0%; Pred. No. 0.28;
Matches 126; Conservative 76; Mismatches 242; Indels 219; Gaps 22;

QY  36 INSTGPGSRALFTPVRSNMDSGDNRASDVGLPVPNMRLLAASEITLNDGFEVLHDHGP 95
Db  395 INVSSSLGQHPAGVITREHVT-----SSASGSSASPSRYSRSATAT-GASITAGSAL 447
QY  96 LDTLNRQ-----IGSSVFRVETQEDGKHIAVGORNGVETSVVLSDOEYARLQSID 145
Db  448 ASANAQKHQSSAAPSSGSSSSRRSSQNDAAATAAG-----GTVVMS----- 490
QY  146 PEGKQKFVETGGRGAGAHAMVTVASDITEARQRIELLEPKGTGSKGAGESKGVGELRE 205
Db  491 -----GTRHGGVQMQAQT-SRQATISLLQPPSKPSSNTTQIAQIPPLFN 535
QY  206 SNSGAENTTETQTSTSTSLR--SDPKLWLALGT-----VATGLIG----- 244
Db  536 RNSTA--TSSAAQPSGTGITTRKIADPKGRIPLNSTAVQGHRTATGAVAAANNGGIPSHRDH 594
QY  245 -----LAATGIVQALALTPEDSPPTTDPDAAASATETATRDOLTKEAFO----- 289
Db  595 AQQQYMNQLTSSWTMSKLINTPAAGGTAATSSSSSSATSTA---PQKSGSQISHAP 651
QY  290 -----NPDNQKYNIDELGNAIPS-----GVLKDDVYANI---EQAKAAGEEA 329
Db  652 TEPVIREDDDENNSENQNGNVPILGGVGQPTSPAVQVPTEDATSSSKKEQQQKASSTP 711
QY  330 KQ-----QAIENNAQAQKYDEQQAQKEELKVSSGAGYGLSGALILGGGIGVAVTAAL 383
Db  712 KESNPVWQNLHLNLSLLKSLDSSAATSYETPRRPGIAG----- 750
QY  384 HRKNQVEQTTTTTTTTTTTTSARTVENKFNNTP-----AQ 419
Db  751 -RRSEPSAATPRRHQTMVDARHLQTPDTPRPHFEDTILDRQMRALYVSTASSRMTR 809
QY  420 GNVDTPGSEDTWESR-----RSMASSTSTFTDTSIGTV----- 454
Db  810 GVLPTPTSNSTSSSFIVEPLTHVAAASPDITTTPTKSTVTSTPYFRRTPSFRMLIVL 869
QY  455 -----QNPYADVKTSLHDSQVTSNSNTSVQNMGNTDSVVYSTIOHPPRD 499
Db  870 LLCNDGLRLWPMIHQSP-----SMPPSQMTAMESLKSQGTGFTVATGGPPQRA 923
QY  500 TTNDNGARLLGNPSAGIQSTYARLALSGRLHDWGGI---TGGS-----NSAVNTSN 547
Db  924 TSQQMSRSATTNSANMGSSGGAASATNQLSGAPSTGASSQQYHPKAPSSSSSST 983
QY  548 NPP 550
Db  984 NPP 986
```

Search completed: May 13, 2005, 11:48:24  
Job time : 46 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:49:17 ; Search time 178 Seconds  
(without alignments)  
1605.282 Million cell updates/sec

Title: US-09-189-415B-11

Perfect score: 558

Sequence: 1 MPIGNLGHNPVNSIPPAP.....SNSAVNTSNPPAGSHRFV 558

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 8

Total number of hits satisfying chosen parameters: 1770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	558	100.0	558	2	Q7DB77 escherichia
2	558	100.0	558	2	Q9R396 escherichia
3	340	60.9	558	2	O85506 escherichia
4	28	5.0	547	2	Q7BHL5 citrobacter
5	28	5.0	547	2	Q9ET11 citrobacter
6	28	5.0	547	2	Q9WKK1 escherichia
7	22	3.9	538	2	O85508 escherichia
8	22	3.9	538	2	Q47014 escherichia
9	22	3.9	538	2	Q47016 escherichia
10	20	3.6	367	2	Q79C12 escherichia
11	14	2.5	549	2	O50190 escherichia
12	14	2.5	550	2	O52147 escherichia
13	14	2.5	551	2	O68258 escherichia
14	14	2.5	552	2	Q9KWH9 escherichia
15	13	2.3	163	2	Q6GED0 escherichia
16	13	2.3	166	2	Q7A060 staphylococ
17	13	2.3	166	2	Q7A2K7 staphylococ
18	13	2.3	166	2	Q7A418 staphylococ
19	13	2.3	166	2	Q6G718 staphylococ
20	13	2.3	271	2	Q94491 dictyosteli
21	13	2.3	498	2	O86J21 dictyosteli
22	12	2.2	56	2	Q8T1V8 dictyosteli
23	12	2.2	89	2	Q9NIP9 trypanosoma
24	12	2.2	106	2	Q6WB03 trypanosoma
25	12	2.2	106	2	Q6WB04 trypanosoma
26	12	2.2	106	2	Q6WB05 trypanosoma
27	12	2.2	107	2	Q6WB01 trypanosoma
28	12	2.2	107	2	Q6WB08 trypanosoma
29	12	2.2	107	2	Q9NIQ1 trypanosoma
30	12	2.2	108	2	Q6WB02 trypanosoma
31	12	2.2	119	2	Q9N6G3 trypanosoma

32	12	2.2	120	2	Q9NIQ0 trypanosoma
33	12	2.2	130	2	Q6WB06 trypanosoma
34	12	2.2	132	2	Q9NIQ3 trypanosoma
35	12	2.2	139	2	Q9C7Q5 arabidopsis
36	12	2.2	150	2	Q8LEL8 arabidopsis
37	12	2.2	200	2	Q9XX76 caenorhabdi
38	12	2.2	291	2	Q94467 dictyosteli
39	12	2.2	336	1	RT09 CANAL
40	12	2.2	337	2	Q86K30 dictyosteli
41	12	2.2	356	2	Q7PZ21 anopheles g
42	12	2.2	364	2	Q9W2Z7 drosophila
43	12	2.2	402	2	O44626 caenorhabdi
44	12	2.2	403	2	Q9XJ33 cyanidium c
45	12	2.2	444	2	Q7S050 neurospora
46	12	2.2	445	2	Q8K3T6 mus musculu
47	12	2.2	458	2	Q9NA83 caenorhabdi
48	12	2.2	486	2	Q7S0Z3 neurospora
49	12	2.2	524	1	HME2 SCHMA
50	12	2.2	540	2	Q7SJM6 dictyosteli
51	12	2.2	569	2	Q83151 murid herpe
52	12	2.2	569	2	Q83183 murid herpe
53	12	2.2	646	2	Q8MNM4 dictyosteli
54	12	2.2	648	2	Q86A81 d similar t
55	12	2.2	649	2	Q8MUF9 dictyosteli
56	12	2.2	659	2	Q94A11 arabidopsis
57	12	2.2	674	2	Q7YZH9 monosiga br
58	12	2.2	693	2	Q96XH6 sulfobobus
59	12	2.2	702	2	Q9V615 drosophila
60	12	2.2	711	2	Q7VYX9 cryptospori
61	12	2.2	717	2	Q8U1H5 pyrococcus
62	12	2.2	718	2	Q9C6C5 arabidopsis
63	12	2.2	722	2	Q9C700 arabidopsis
64	12	2.2	725	2	Q8T124 dictyosteli
65	12	2.2	735	2	Q86LA3 dictyosteli
66	12	2.2	766	2	Q6Q7J0 drosophila
67	12	2.2	776	2	Q7Q5E1 anopheles g
68	12	2.2	798	2	Q86K66 dictyosteli
69	12	2.2	802	1	XYND_RUMFL
70	12	2.2	802	2	Q9S310 ruminococcu
71	12	2.2	832	2	Q86H63 dictyosteli
72	12	2.2	855	2	Q869V0 dictyosteli
73	12	2.2	860	2	Q23916 dictyosteli
74	12	2.2	872	2	Q26257 dictyosteli
75	12	2.2	986	2	Q9N9W3 haemochus
76	12	2.2	1026	1	GLT_DROME
77	12	2.2	1026	2	Q9VQZ3 drosophila
78	12	2.2	1123	2	Q8U181 dictyosteli
79	12	2.2	1247	2	Q9U9T1 dictyosteli
80	12	2.2	1247	2	Q9Y0T2 dictyosteli
81	12	2.2	1256	2	Q875X6 oikopleura
82	12	2.2	1258	2	Q86KB3 dictyosteli
83	12	2.2	1297	2	Q8SSS5 dictyosteli
84	12	2.2	1313	2	Q86IU8 dictyosteli
85	12	2.2	1326	2	Q86I09 dictyosteli
86	12	2.2	1387	2	Q86H9K1 dictyosteli
87	12	2.2	1438	2	Q86HL0 dictyosteli
88	12	2.2	1461	2	Q86AZ9 dictyosteli
89	12	2.2	1565	2	Q86K18 dictyosteli
90	12	2.2	1695	2	Q9NK53 drosophila
91	12	2.2	1695	2	Q9U9R5 drosophila
92	12	2.2	1704	1	VILD_DICDI
93	12	2.2	1711	2	Q9U9R4 drosophila
94	12	2.2	1711	2	Q9VJL0 drosophila
95	12	2.2	1838	2	Q86IV6 dictyosteli
96	12	2.2	1845	2	Q86JH8 dictyosteli
97	12	2.2	1909	2	Q86IF3 dictyosteli
98	12	2.2	1929	2	Q8IHL2 dictyosteli
99	12	2.2	2159	2	Q86IU5 dictyosteli
100	12	2.2	2454	2	Q86B02 dictyosteli
101	12	2.2	3432	2	Q95PH9 dictyosteli
102	12	2.2	3712	1	LMA_DROME
103	12	2.2	3712	2	Q9VFW0 drosophila
104	12	2.2	3770	2	Q869R6 dictyosteli

105	12	2.2	4377	1	ANK3_HUMAN	Q12955 homo sapien	178	11	2.0	165	2	Q94671	Q94671 plasmodium
106	12	2.2	4493	2	Q861H5	Q861h5 dictyosteli	179	11	2.0	166	2	Q7PGC6	Q7pgc6 anopheles g
107	11	2.0	56	2	Q01601	Q01601 pneumocysti	180	11	2.0	172	2	Q9vzb8	Q9vzb8 drosophila
108	11	2.0	56	2	Q861E6	Q861e6 dictyosteli	181	11	2.0	179	2	Q9XZ40	Q9XZ40 plasmodium
109	11	2.0	58	2	Q6TUI3	Q6tui3 rattus norv	182	11	2.0	183	1	AAC1_DICDI	Pl4195 dictyosteli
110	11	2.0	64	2	Q9N114	Q9n114 plasmodium	183	11	2.0	183	2	Q6UA37	Q6ua97 plasmodium
111	11	2.0	67	2	Q95UY4	Q95uy4 plasmodium	184	11	2.0	183	2	Q9U0B5	Q9u0b5 plasmodium
112	11	2.0	67	2	Q95UY6	Q95uy6 plasmodium	185	11	2.0	183	2	Q9U0B6	Q9u0b6 plasmodium
113	11	2.0	67	2	Q86JN9	Q86jn9 dictyosteli	186	11	2.0	184	2	Q9TY34	Q9ty34 plasmodium
114	11	2.0	71	2	Q962U5	Q962u5 plasmodium	187	11	2.0	184	2	Q9U0B8	Q9u0b8 plasmodium
115	11	2.0	71	2	Q9N103	Q9n103 plasmodium	188	11	2.0	184	2	Q9U0B9	Q9u0b9 plasmodium
116	11	2.0	71	2	Q6R5F0	Q6r5f0 mus musculu	189	11	2.0	191	2	Q9U0C1	Q9u0c1 plasmodium
117	11	2.0	72	2	Q9W3Q9	Q9w3q9 drosophila	190	11	2.0	195	2	Q93948	Q93948 candida alb
118	11	2.0	74	2	Q61023	Q61023 trypanosoma	191	11	2.0	196	2	Q25947	Q25947 plasmodium
119	11	2.0	74	2	Q9N121	Q9n121 plasmodium	192	11	2.0	196	2	Q25951	Q25951 plasmodium
120	11	2.0	78	2	Q962U6	Q962u6 plasmodium	193	11	2.0	202	2	Q01615	Q01615 pneumocysti
121	11	2.0	86	2	Q9TVF2	Q9tvf2 trypanosoma	194	11	2.0	202	2	Q25952	Q25952 plasmodium
122	11	2.0	86	2	Q61058	Q61058 trypanosoma	195	11	2.0	203	2	Q015806	Q015806 plasmodium
123	11	2.0	98	2	Q9N115	Q9n115 plasmodium	196	11	2.0	203	2	Q9U0C0	Q9u0c0 plasmodium
124	11	2.0	102	2	Q61033	Q61033 trypanosoma	197	11	2.0	203	2	Q9W2W1	Q9w2w1 drosophila
125	11	2.0	107	2	Q61050	Q61050 trypanosoma	198	11	2.0	203	2	Q9LVA3	Q9lva3 arabidopsis
126	11	2.0	108	2	Q9XWNO	Q9xwn0 caenorhabdi	199	11	2.0	204	2	Q8H7E2	Q8h7e2 arabidopsis
127	11	2.0	109	2	Q01619	Q01619 pneumocysti	200	11	2.0	205	2	Q15777	Q15777 trypanosoma
128	11	2.0	109	2	Q6PN90	Q6pn90 plasmodium	201	11	2.0	205	2	Q15911	Q15911 dictyosteli
129	11	2.0	110	2	Q9N116	Q9n116 plasmodium	202	11	2.0	206	2	Q9WFZ9	Q9wzfz9 amsacta alb
130	11	2.0	110	2	Q9N117	Q9n117 plasmodium	203	11	2.0	207	2	Q25701	Q25701 plasmodium
131	11	2.0	110	2	Q9N118	Q9n118 plasmodium	204	11	2.0	208	2	Q25949	Q25949 plasmodium
132	11	2.0	111	2	Q9N126	Q9n126 plasmodium	205	11	2.0	209	2	Q61055	Q61055 trypanosoma
133	11	2.0	112	2	Q6PPZ9	Q6ppz9 plasmodium	206	11	2.0	210	2	Q9Y0Z5	Q9y0z5 trypanosoma
134	11	2.0	113	2	Q9BJN8	Q9bjn8 plasmodium	207	11	2.0	216	2	Q962W6	Q962w6 trypanosoma
135	11	2.0	115	2	Q61046	Q61046 trypanosoma	208	11	2.0	217	1	SGS3_DROSI	Pl3729 drosophila
136	11	2.0	115	2	Q9B7Q9	Q9bjq9 plasmodium	209	11	2.0	220	2	Q9U0E3	Q9u0e3 plasmodium
137	11	2.0	116	2	Q9B7Q2	Q9bjq2 plasmodium	210	11	2.0	229	2	Q9VIA7	Q9via7 drosophila
138	11	2.0	118	2	Q9NIP8	Q9nip8 trypanosoma	211	11	2.0	236	2	Q7SDS9	Q7sds9 neurospora
139	11	2.0	119	2	Q61034	Q61034 trypanosoma	212	11	2.0	239	2	Q9W2R5	Q9w2r5 drosophila
140	11	2.0	119	2	Q9BJP7	Q9bjp7 plasmodium	213	11	2.0	242	2	Q9VDN0	Q9vdsn0 drosophila
141	11	2.0	120	2	Q9GQY0	Q9gqy0 plasmodium	214	11	2.0	244	2	Q9U2Q6	Q9u2q6 caenorhabdi
142	11	2.0	121	2	Q6WAZ9	Q6waz9 trypanosoma	215	11	2.0	245	2	Q9XWP2	Q9xwp2 caenorhabdi
143	11	2.0	122	2	Q15774	Q15774 trypanosoma	216	11	2.0	246	2	Q76P06	Q76p06 dictyosteli
144	11	2.0	122	2	Q9S7P8	Q9s7p8 arabidopsis	217	11	2.0	259	2	Q86IM4	Q86im4 dictyosteli
145	11	2.0	123	2	Q61027	Q61027 trypanosoma	218	11	2.0	260	2	Q8IT83	Q8it83 plasmodium
146	11	2.0	123	2	P90602	P90602 trypanosoma	219	11	2.0	264	2	Q96920	Q96920 plasmodium
147	11	2.0	124	2	Q6PQ00	Q6pq00 plasmodium	220	11	2.0	274	1	MSA2_PLAF6	P50497 plasmodium
148	11	2.0	125	2	Q61025	Q61025 trypanosoma	221	11	2.0	274	2	Q7SC14	Q7sc14 neurospora
149	11	2.0	125	2	Q962W4	Q962w4 trypanosoma	222	11	2.0	277	2	Q86IC7	Q86ic7 dictyosteli
150	11	2.0	126	2	Q61021	Q61021 trypanosoma	223	11	2.0	278	2	Q25862	Q25862 plasmodium
151	11	2.0	126	2	Q61056	Q61056 trypanosoma	224	11	2.0	280	2	Q17639	Q17639 caenorhabdi
152	11	2.0	128	2	P90603	P90603 trypanosoma	225	11	2.0	283	2	Q86IL5	Q86il5 dictyosteli
153	11	2.0	128	2	Q9N1Q2	Q9niq2 trypanosoma	226	11	2.0	284	2	Q20202	Q20202 caenorhabdi
154	11	2.0	128	2	Q49334	Q49334 mycoplasma	227	11	2.0	293	2	Q25785	Q25785 plasmodium
155	11	2.0	131	2	Q9D9N0	Q9d9n0 mus musculu	228	11	2.0	295	2	Q9QYL6	Q9qyl6 mus musculu
156	11	2.0	136	2	Q86KRO	Q86kro dictyosteli	229	11	2.0	295	2	Q9Z2H8	Q9z2h8 mus musculu
157	11	2.0	139	2	Q61037	Q61037 trypanosoma	230	11	2.0	302	2	Q95U23	Q95u23 drosophila
158	11	2.0	139	2	P90601	P90601 trypanosoma	231	11	2.0	303	2	Q9ILM9	Q9ilm9 white spot
159	11	2.0	139	2	Q6WAZ8	Q6waz8 trypanosoma	232	11	2.0	304	1	YOOB_CABEL	Q09300 caenorhabdi
160	11	2.0	139	2	Q7Z0N2	Q7z0n2 caenorhabdi	233	11	2.0	306	2	Q9QV14	Q9qvl4 mus musculu
161	11	2.0	140	2	Q962W5	Q962w5 trypanosoma	234	11	2.0	312	2	Q01824	Q01824 pneumocysti
162	11	2.0	143	2	Q15776	Q15776 trypanosoma	235	11	2.0	312	2	Q6VZS5	Q6vzs5 canarypox v
163	11	2.0	143	2	Q8IT82	Q8it82 plasmodium	236	11	2.0	320	2	Q9U448	Q9u448 dictyosteli
164	11	2.0	146	2	Q6RUB8	Q6rub8 chimpanzee	237	11	2.0	322	2	Q8SXG9	Q8sxg9 drosophila
165	11	2.0	147	2	Q6RUE3	Q6rue3 chimpanzee	238	11	2.0	323	2	Q86AP2	Q86ap2 dictyosteli
166	11	2.0	147	2	Q6RUE8	Q6rue8 chimpanzee	239	11	2.0	327	2	Q25334	Q25334 leishmania
167	11	2.0	148	2	Q61019	Q61019 trypanosoma	240	11	2.0	335	2	Q96U23	Q96u23 neurospora
168	11	2.0	148	2	Q6WB00	Q6wb00 trypanosoma	241	11	2.0	335	2	Q86HM2	Q86hm2 dictyosteli
169	11	2.0	148	2	Q6RUC1	Q6ruc1 chimpanzee	242	11	2.0	341	2	Q8IMS9	Q8ims9 drosophila
170	11	2.0	150	2	Q9BJQ6	Q9bjq6 plasmodium	243	11	2.0	344	2	Q76741	Q76741 dictyosteli
171	11	2.0	150	2	Q9B7Q7	Q9bjq7 plasmodium	244	11	2.0	350	2	Q7Q1R0	Q7q1r0 anopheles g
172	11	2.0	150	2	Q9GQX3	Q9gqx3 plasmodium	245	11	2.0	362	2	Q7YUQ6	Q7yuq6 trypanosoma
173	11	2.0	150	2	Q9VTF1	Q9vtf1 drosophila	246	11	2.0	364	2	Q7S2P4	Q7s2p4 neurospora
174	11	2.0	155	2	Q6U5F5	Q6u5f5 plasmodium	247	11	2.0	364	2	Q7YUQ5	Q7yuq5 trypanosoma
175	11	2.0	157	2	Q25713	Q25713 plasmodium	248	11	2.0	364	2	Q688R2	Q688r2 oryza sativ
176	11	2.0	160	2	Q94669	Q94669 plasmodium	249	11	2.0	365	2	Q7YUW8	Q7yuw8 trypanosoma
177	11	2.0	164	2	Q9B7Q5	Q9bjq5 plasmodium	250	11	2.0	365	2	Q869R5	Q869r5 dictyosteli

251	11	2.0	366	2	Q9YZ40	Q9YZ40 chimpanzee	324	11	2.0	571	2	Q8NJP9	Q8NJP9 candida alb
252	11	2.0	367	2	Q7YUQ8	Q7YUQ8 trypanosoma	325	11	2.0	573	2	Q8TOR9	Q8TOR9 drosophila
253	11	2.0	369	2	Q7YUQ1	Q7YUQ1 trypanosoma	326	11	2.0	574	2	Q9DSX5	Q9DSX5 m mus muscu
254	11	2.0	369	2	Q7YUQ2	Q7YUQ2 trypanosoma	327	11	2.0	579	2	Q7SDE0	Q7SDE0 neurospora
255	11	2.0	369	2	Q7YUQ3	Q7YUQ3 trypanosoma	328	11	2.0	600	2	Q17490	Q17490 anopheles g
256	11	2.0	369	2	Q7YUQ4	Q7YUQ4 trypanosoma	329	11	2.0	600	2	Q7PV58	Q7PV58 anopheles g
257	11	2.0	374	2	Q7QCS5	Q7QCS5 anopheles g	330	11	2.0	602	2	Q86JM6	Q86JM6 dictyosteli
258	11	2.0	385	2	Q7LZ8R	Q7LZ8R ictalurus p	331	11	2.0	627	2	Q6UN14	Q6UN14 leishmania
259	11	2.0	386	2	Q01759	Q01759 pneumocysti	332	11	2.0	636	2	Q8MP77	Q8MP77 dictyosteli
260	11	2.0	386	2	Q9GZ9H	Q9GZ9H caenorhabdi	333	11	2.0	648	2	Q95QX0	Q95QX0 caenorhabdi
261	11	2.0	390	2	Q8MP08	Q8MP08 bombyx mori	334	11	2.0	657	2	Q86J19	Q86J19 dictyosteli
262	11	2.0	392	2	Q8IIC1	Q8IIC1 plasmodium	335	11	2.0	660	2	Q94485	Q94485 dictyosteli
263	11	2.0	392	2	Q69Z58	Q69Z58 mus musculu	336	11	2.0	667	2	Q7VYV0	Q7VYV0 cryptospori
264	11	2.0	398	2	Q22902	Q22902 caenorhabdi	337	11	2.0	680	2	Q93374	Q93374 caenorhabdi
265	11	2.0	399	2	Q86TX1	Q86TX1 dictyosteli	338	11	2.0	681	2	Q9XUS9	Q9XUS9 caenorhabdi
266	11	2.0	404	2	Q86SB6	Q86SB6 drosophila	339	11	2.0	684	2	Q8T848	Q8T848 dictyosteli
267	11	2.0	417	2	Q01760	Q01760 pneumocysti	340	11	2.0	689	2	Q9Y1I4	Q9Y1I4 haemochus
268	11	2.0	422	2	Q91BH8	Q91BH8 spodoptera	341	11	2.0	693	2	Q07241	Q07241 pneumocysti
269	11	2.0	425	2	Q15755	Q15755 dictyosteli	342	11	2.0	697	1	CRAC_DICDI	CRAC_DICDI
270	11	2.0	425	2	Q86AP8	Q86AP8 dictyosteli	343	11	2.0	701	2	Q9U233	Q9U233 caenorhabdi
271	11	2.0	425	2	Q9GZFS	Q9GZFS caenorhabdi	344	11	2.0	714	2	Q9NGW9	Q9NGW9 dictyosteli
272	11	2.0	426	2	Q961P7	Q961P7 drosophila	345	11	2.0	730	2	Q7QCW4	Q7QCW4 anopheles g
273	11	2.0	426	2	Q9EN59	Q9EN59 spodoptera	346	11	2.0	732	2	Q95Q40	Q95Q40 caenorhabdi
274	11	2.0	427	2	Q86I17	Q86I17 dictyosteli	347	11	2.0	732	2	Q86L25	Q86L25 dictyosteli
275	11	2.0	433	2	Q9VFD1	Q9VFD1 drosophila	348	11	2.0	734	2	Q86AK8	Q86AK8 dictyosteli
276	11	2.0	438	2	Q86TC6	Q86TC6 dictyosteli	349	11	2.0	737	2	Q86IX1	Q86IX1 dictyosteli
277	11	2.0	440	2	Q20659	Q20659 caenorhabdi	350	11	2.0	740	2	Q8T2A4	Q8T2A4 dictyosteli
278	11	2.0	442	2	Q9BY67	Q9BY67 homo sapien	351	11	2.0	742	2	Q23766	Q23766 chromomus
279	11	2.0	445	2	Q7Q956	Q7Q956 anopheles g	352	11	2.0	746	2	Q9V515	Q9V515 drosophila
280	11	2.0	448	2	Q8R4L1	Q8R4L1 mus musculu	353	11	2.0	748	2	Q44416	Q44416 chironomus
281	11	2.0	448	2	Q9N3I0	Q9N3I0 caenorhabdi	354	11	2.0	754	2	Q6JXF2	Q6JXF2 schistosoma
282	11	2.0	453	2	Q870S8	Q870S8 neurospora	355	11	2.0	756	2	Q86IG6	Q86IG6 dictyosteli
283	11	2.0	456	2	Q8R5M8	Q8R5M8 mus musculu	356	11	2.0	758	2	Q8MPP3	Q8MPP3 drosophila
284	11	2.0	465	2	Q86KH4	Q86KH4 dictyosteli	357	11	2.0	761	2	Q9VD08	Q9VD08 drosophila
285	11	2.0	467	1	KG3H_DICDI	KG3H_DICDI	358	11	2.0	769	2	Q17921	Q17921 caenorhabdi
286	11	2.0	468	2	Q55279	Q55279 avian adeno	359	11	2.0	771	2	Q86H13	Q86H13 dictyosteli
287	11	2.0	474	2	Q7S290	Q7S290 neurospora	360	11	2.0	781	2	Q52753	Q52753 ruminococcu
288	11	2.0	476	2	Q6AYP5	Q6AYP5 rattus norv	361	11	2.0	782	2	Q75JA4	Q75JA4 dictyosteli
289	11	2.0	477	2	Q23993	Q23993 drosophila	362	11	2.0	793	1	REGA_DICDI	REGA_DICDI
290	11	2.0	482	2	Q7SE49	Q7SE49 neurospora	363	11	2.0	795	2	Q9Y1J5	Q9Y1J5 dictyosteli
291	11	2.0	484	1	OAR1_LOCMI	OAR1_LOCMI	364	11	2.0	813	2	P90538	P90538 dictyosteli
292	11	2.0	484	1	OAR2_LOCMI	OAR2_LOCMI	365	11	2.0	814	2	Q7S1S5	Q7S1S5 neurospora
293	11	2.0	493	2	Q7S3X3	Q7S3X3 neurospora	366	11	2.0	823	2	Q75JE3	Q75JE3 dictyosteli
294	11	2.0	503	2	Q13269	Q13269 caenorhabdi	367	11	2.0	831	2	Q15756	Q15756 dictyosteli
295	11	2.0	504	2	Q6CCD0	Q6CCD0 yarrowia li	368	11	2.0	837	2	Q7QFG1	Q7QFG1 anopheles g
296	11	2.0	511	2	Q6CCD8	Q6CCD8 dictyosteli	369	11	2.0	839	2	Q8T1P1	Q8T1P1 dictyosteli
297	11	2.0	512	1	WR33_ARATH	WR33_ARATH	370	11	2.0	844	2	Q86AP1	Q86AP1 brachydanio
298	11	2.0	513	2	Q6V4A2	Q6V4A2 drosophila	371	11	2.0	867	2	Q90X49	Q90X49 dictyosteli
299	11	2.0	513	2	Q86P22	Q86P22 drosophila	372	11	2.0	868	2	Q86IC5	Q86IC5 dictyosteli
300	11	2.0	513	2	Q9W3K2	Q9W3K2 drosophila	373	11	2.0	872	2	P90523	P90523 dictyosteli
301	11	2.0	517	1	LA1C_DIACA	LA1C_DIACA	374	11	2.0	874	2	Q76535	Q76535 dictyosteli
302	11	2.0	518	2	Q43753	Q43753 dianthus ca	375	11	2.0	887	2	Q23913	Q23913 dictyosteli
303	11	2.0	519	2	Q09592	Q09592 caenorhabdi	376	11	2.0	889	2	Q23895	Q23895 dictyosteli
304	11	2.0	520	2	Q87LE3	Q87LE3 neurospora	377	11	2.0	893	2	Q8T7Y3	Q8T7Y3 dictyosteli
305	11	2.0	520	2	Q9VHU5	Q9VHU5 drosophila	378	11	2.0	895	2	Q7KWT7	Q7KWT7 dictyosteli
306	11	2.0	522	2	Q75JX8	Q75JX8 dictyosteli	379	11	2.0	895	2	Q86A69	Q86A69 dictyosteli
307	11	2.0	524	2	Q653F4	Q653F4 oryza sativ	380	11	2.0	899	2	Q7S5JH3	Q7S5JH3 dictyosteli
308	11	2.0	525	1	NU62_RAT	NU62_RAT	381	11	2.0	910	2	Q7S602	Q7S602 neurospora
309	11	2.0	526	1	NU62_MOUSE	NU62_MOUSE	382	11	2.0	911	2	Q7KWL1	Q7KWL1 dictyosteli
310	11	2.0	526	2	Q8VAG9	Q8VAG9 mus musculu	383	11	2.0	919	2	Q86J16	Q86J16 dictyosteli
311	11	2.0	527	2	Q8QTA8	Q8QTA8 white spot	384	11	2.0	930	2	Q8T1S8	Q8T1S8 dictyosteli
312	11	2.0	530	2	Q61209	Q61209 caenorhabdi	385	11	2.0	931	2	Q86K99	Q86K99 dictyosteli
313	11	2.0	534	2	Q7KTI5	Q7KTI5 drosophila	386	11	2.0	932	2	Q15924	Q15924 dictyosteli
314	11	2.0	538	2	Q75JS0	Q75JS0 dictyosteli	387	11	2.0	937	2	Q86L47	Q86L47 dictyosteli
315	11	2.0	542	2	Q966U0	Q966U0 dictyosteli	388	11	2.0	947	2	Q86H44	Q86H44 dictyosteli
316	11	2.0	543	2	Q84UZ6	Q84UZ6 mesembryant	389	11	2.0	956	2	Q00908	Q00908 cryptospori
317	11	2.0	547	2	Q7PDW3	Q7PDW3 anopheles g	390	11	2.0	961	2	Q8IC12	Q8IC12 plasmodium
318	11	2.0	548	2	Q7S905	Q7S905 neurospora	391	11	2.0	967	2	Q86AS0	Q86AS0 dictyosteli
319	11	2.0	551	1	PPBJ_RAT	PPBJ_RAT	392	11	2.0	975	2	Q15743	Q15743 dictyosteli
320	11	2.0	552	2	Q19659	Q19659 caenorhabdi	393	11	2.0	984	2	Q9Y1P7	Q9Y1P7 cryptospori
321	11	2.0	559	1	PPBI_MOUSE	PPBI_MOUSE	394	11	2.0	992	2	Q86H50	Q86H50 dictyosteli
322	11	2.0	559	2	Q9VN36	Q9VN36 mus musculu	395	11	2.0	995	2	Q95WL0	Q95WL0 lymantria d
323	11	2.0	562	2	Q9JKS8	Q9JKS8 rattus norv	396	11	2.0	1000	2	Q8ITD4	Q8ITD4 helicoverpa

397	11	2.0	1002	2	074674	074674 pneumocysti	470	11	2.0	1598	2	Q95YM8	Q95YM8 apis mellif
398	11	2.0	1010	2	086J70	086j70 dictyosteli	471	11	2.0	1635	2	017412	017412 aedes aegyp
399	11	2.0	1010	2	09U6W2	09u6w2 heliothis v	472	11	2.0	1641	2	02GR23	02gr23 caenorhabdi
400	11	2.0	1011	2	09NHZ9	09nhz9 helicoverpa	473	11	2.0	1670	2	023901	023901 dictyosteli
401	11	2.0	1013	2	08WSZ2	08wsz2 helicoverpa	474	11	2.0	1671	2	0869S5	0869s5 dictyosteli
402	11	2.0	1014	2	0962B4	0962b4 helicoverpa	475	11	2.0	1672	2	0869N0	0869n0 dictyosteli
403	11	2.0	1014	2	06UY54	06uy54 helicoverpa	476	11	2.0	1679	2	07KT16	07kt16 drosophila
404	11	2.0	1014	2	06UY55	06uy55 helicoverpa	477	11	2.0	1682	2	015054	015054 homo sapien
405	11	2.0	1015	2	08FAG0	08fag0 dictyosteli	478	11	2.0	1693	2	086H41	086h41 dictyosteli
406	11	2.0	1024	2	08FAG9	08fag9 dictyosteli	479	11	2.0	1723	2	086H41	086h41 dictyosteli
407	11	2.0	1028	2	0869U1	0869u1 dictyosteli	480	11	2.0	1728	2	08SSU4	08ssu4 dictyosteli
408	11	2.0	1037	2	08IS18	08is18 dictyosteli	481	11	2.0	1731	2	09TW28	09tw28 dictyosteli
409	11	2.0	1051	2	086142	086142 dictyosteli	482	11	2.0	1787	2	086KX8	086kx8 dictyosteli
410	11	2.0	1056	2	09W294	09w294 drosophila	483	11	2.0	1804	2	08WPD7	08wpd7 ciona intes
411	11	2.0	1065	2	001828	001828 pneumocysti	484	11	2.0	1832	2	086S03	086s03 cryptospori
412	11	2.0	1069	2	096V12	096v12 pneumocysti	485	11	2.0	1849	2	086LAL	086lal dictyosteli
413	11	2.0	1069	2	06AHV5	06ahv5 pneumocysti	486	11	2.0	1853	2	07KT96	07kt96 drosophila
414	11	2.0	1069	2	09XYP1	09xyp1 dictyosteli	487	11	2.0	1858	1	P3K2_DICDI	P3k2_DICDI
415	11	2.0	1070	2	078734	078734 pneumocysti	488	11	2.0	1873	2	06S003	06s003 dictyosteli
416	11	2.0	1071	2	096V10	096v10 pneumocysti	489	11	2.0	1893	2	09NKC9	09nkc9 drosophila
417	11	2.0	1076	2	001830	001830 pneumocysti	490	11	2.0	1955	2	086K52	086k52 dictyosteli
418	11	2.0	1077	2	06AHS1	06ahs1 pneumocysti	491	11	2.0	2015	2	09U5Y1	09u5y1 dictyosteli
419	11	2.0	1079	2	096V11	096v11 pneumocysti	492	11	2.0	2048	2	086JW3	086jw3 dictyosteli
420	11	2.0	1081	2	06AHT3	06ahnt3 pneumocysti	493	11	2.0	2084	2	0869U3	0869u3 dictyosteli
421	11	2.0	1083	2	012075	012075 pneumocysti	494	11	2.0	2102	2	0861W7	0861w7 d similar t
422	11	2.0	1084	2	086LA6	086la6 dictyosteli	495	11	2.0	2107	2	086M87	086m87 dictyosteli
423	11	2.0	1088	2	09VPC7	09vpc7 drosophila	496	11	2.0	2183	2	07KMS4	07kms4 dictyosteli
424	11	2.0	1089	2	044007	044007 dictyosteli	497	11	2.0	2208	2	086H44	086h44 dictyosteli
425	11	2.0	1092	2	09UVY2	09uvy2 pneumocysti	498	11	2.0	2284	2	086166	086166 dictyosteli
426	11	2.0	1093	1	P14K_DICDI	P14k_DICDI	499	11	2.0	2296	2	09VNE0	09vne0 drosophila
427	11	2.0	1099	1	PLC1_CANAL	PLC1_CANAL	500	11	2.0	2344	2	086H67	086h67 dictyosteli
428	11	2.0	1123	2	086K40	086k40 dictyosteli	501	11	2.0	2528	2	086J36	086j36 dictyosteli
429	11	2.0	1131	2	07S0U4	07s0u4 neurospora	502	11	2.0	2536	2	0817C4	0817c4 dictyosteli
430	11	2.0	1136	2	09V5A6	09v5a6 drosophila	503	11	2.0	2621	2	08SSW5	08ssw5 dictyosteli
431	11	2.0	1138	2	08IS12	08is12 dictyosteli	504	11	2.0	2625	2	08MMZ9	08mmz9 dictyosteli
432	11	2.0	1155	2	07S552	07s552 neurospora	505	11	2.0	2650	2	086AF1	086af1 dictyosteli
433	11	2.0	1158	2	08IU40	08iu40 dictyosteli	506	11	2.0	2678	2	09NDS4	09nds4 dictyosteli
434	11	2.0	1163	2	0869M3	0869m3 dictyosteli	507	11	2.0	2722	2	086JN0	086jn0 dictyosteli
435	11	2.0	1163	2	0861J3	0861j3 dictyosteli	508	11	2.0	2843	2	0963L5	0963l5 dictyosteli
436	11	2.0	1166	2	081P52	081p52 drosophila	509	11	2.0	2994	2	0952G5	0952g5 dictyosteli
437	11	2.0	1192	2	0869E1	0869e1 dictyosteli	510	11	2.0	3177	2	09N1A3	09n1a3 dictyosteli
438	11	2.0	1212	2	086HH0	086hh0 dictyosteli	511	11	2.0	3295	2	066GT3	066gt3 rattus norv
439	11	2.0	1222	2	086HL6	086hl6 dictyosteli	512	11	2.0	3446	2	086AC8	086ac8 dictyosteli
440	11	2.0	1222	2	086SA4	086sa4 dictyosteli	513	11	2.0	3535	2	081C29	081c29 plasmodium
441	11	2.0	1224	2	07Q0G7	07q0g7 anopheles g	514	11	2.0	3550	2	066GT4	066gt4 rattus norv
442	11	2.0	1227	2	06T872	06t872 dictyosteli	515	11	2.0	6118	2	081396	081396 plasmodium
443	11	2.0	1240	2	09DWH8	09dwh8 rat cytomeg	516	11	2.0	23015	2	081Q18	081q18 drosophila
444	11	2.0	1264	2	086H34	086h34 dictyosteli	517	10	1.8	46	2	08J4I7	08j4i7 chimpanzee
445	11	2.0	1265	2	059920	059920 pneumocysti	518	10	1.8	86	2	026881	026881 trypanosoma
446	11	2.0	1271	1	Y338_MYCGE	P47580 mycoplasma	519	10	1.8	86	2	07PFH5	07pfh5 anopheles g
447	11	2.0	1271	2	086123	086123 dictyosteli	520	10	1.8	113	2	061052	061052 trypanosoma
448	11	2.0	1285	2	07SGA3	07sga3 neurospora	521	10	1.8	120	2	09N1Q7	09n1q7 trypanosoma
449	11	2.0	1298	2	086G47	086g47 dictyosteli	522	10	1.8	125	2	096VJ3	096vj3 pneumocysti
450	11	2.0	1303	2	066S84	066s84 okopileura	523	10	1.8	127	2	08LBJ3	08lbj3 arabidopsis
451	11	2.0	1307	2	086J7U4	086j7u4 dictyosteli	524	10	1.8	127	2	09LF22	09lf22 arabidopsis
452	11	2.0	1310	2	0861X0	0861x0 dictyosteli	525	10	1.8	134	2	P89278	P89278 xestia C-ni
453	11	2.0	1364	2	08T245	08t245 dictyosteli	526	10	1.8	137	2	08QHP0	08qhp0 human cytom
454	11	2.0	1380	2	086J7U5	086j7u5 dictyosteli	527	10	1.8	137	2	08QVD8	08qvd8 human cytom
455	11	2.0	1398	2	0966T5	0966t5 dictyosteli	528	10	1.8	137	2	08QVD9	08qvd9 human cytom
456	11	2.0	1402	2	096668	096668 dictyosteli	529	10	1.8	138	2	08Q326	08q326 porcine her
457	11	2.0	1404	2	086180	086180 dictyosteli	530	10	1.8	143	2	061028	061028 trypanosoma
458	11	2.0	1405	2	07KWK0	07kwk0 dictyosteli	531	10	1.8	144	2	061KH6	061kh6 drosophila
459	11	2.0	1450	2	086A05	086a05 dictyosteli	532	10	1.8	146	2	06RUD1	06rud1 chimpanzee
460	11	2.0	1483	2	07KWKU3	07kwk3 dictyosteli	533	10	1.8	146	2	06RUD5	06rud5 chimpanzee
461	11	2.0	1485	2	07KWP2	07kwp2 dictyosteli	534	10	1.8	146	2	06RUE0	06rue0 chimpanzee
462	11	2.0	1489	2	06NP07	06np07 drosophila	535	10	1.8	146	2	06RUE5	06rue5 chimpanzee
463	11	2.0	1513	2	07SCN1	07scn1 neurospora	536	10	1.8	147	2	06RUC3	06ruc3 chimpanzee
464	11	2.0	1518	2	086AN8	086an8 dictyosteli	537	10	1.8	147	2	06RUC3	06ruc3 chimpanzee
465	11	2.0	1523	2	086L99	086l99 dictyosteli	538	10	1.8	147	2	06RUD8	06rud8 chimpanzee
466	11	2.0	1524	2	09V854	09v854 drosophila	539	10	1.8	147	2	06RUE9	06rue9 chimpanzee
467	11	2.0	1529	2	081S21	081s21 dictyosteli	540	10	1.8	147	2	06RUF2	06ruf2 chimpanzee
468	11	2.0	1529	2	09QC2	09qc2 dictyosteli	541	10	1.8	148	2	06RUB6	06rub6 chimpanzee
469	11	2.0	1556	2	08ST25	08st25 dictyosteli	542	10	1.8	148	2	06RUE1	06rue1 chimpanzee

543	1.8	152	2	Q7S996	Q7S996 neurospora	616	10	1.8	630	2	Q22225	Q22225 caenorhabdi
544	1.8	157	2	Q7PRF6	Q7PRF6 anopheles g	617	10	1.8	634	2	Q7KX02	Q7KX02 caenorhabdi
545	1.8	165	2	Q26877	Q26877 trypanosoma	618	10	1.8	648	1	KAPC_DICDI	P34099 dictyosteli
546	1.8	188	2	Q9NAQ1	Q9NAQ1 caenorhabdi	619	10	1.8	648	2	Q9TXB8	Q9TXB8 dictyosteli
547	1.8	194	2	Q6IK11	Q6IK11 drosophila	620	10	1.8	650	1	ARE2_SACBA	Q87612 saccharomyc
548	1.8	194	2	Q7PIY5	Q7PIY5 anopheles g	621	10	1.8	650	2	Q7SE54	Q7SE54 neurospora
549	1.8	195	2	Q71630	Q71630 caenorhabdi	622	10	1.8	659	2	Q6ZBH4	Q6ZBH4 oryza sativ
550	1.8	195	2	Q9W4V3	Q9W4V3 drosophila	623	10	1.8	662	1	MUC1_XENLA	Q05049 xenopus lae
551	1.8	197	2	Q26878	Q26878 trypanosoma	624	10	1.8	665	2	Q9N2V0	Q9N2V0 caenorhabdi
552	1.8	201	2	Q6WAZ6	Q6WAZ6 trypanosoma	625	10	1.8	696	2	Q86HT7	Q86HT7 dictyosteli
553	1.8	203	2	Q7S4M2	Q7S4M2 neurospora	626	10	1.8	699	2	Q9MAH1	Q9MAH1 arabidopsis
554	1.8	206	2	Q7NDW9	Q7NDW9 gloeobacter	627	10	1.8	711	2	Q7S623	Q7S623 neurospora
555	1.8	207	2	Q26879	Q26879 trypanosoma	628	10	1.8	732	2	Q76NV3	Q76NV3 dictyosteli
556	1.8	211	2	Q00026	Q00026 ajilomyces	629	10	1.8	734	2	Q869R9	Q869R9 dictyosteli
557	1.8	211	2	Q6WAZ3	Q6WAZ3 trypanosoma	630	10	1.8	741	2	Q9PY56	Q9PY56 xestia c-ni
558	1.8	219	2	Q7S8U5	Q7S8U5 neurospora	631	10	1.8	743	2	Q7PQ19	Q7PQ19 anopheles g
559	1.8	225	2	Q7PK76	Q7PK76 anopheles g	632	10	1.8	752	2	Q8MNE2	Q8MNE2 dictyosteli
560	1.8	227	2	Q7S8L4	Q7S8L4 neurospora	633	10	1.8	756	2	Q44417	Q44417 chironomus
561	1.8	231	2	Q9BIQ7	Q9BIQ7 cooperia pu	634	10	1.8	756	2	Q44418	Q44418 chironomus
562	1.8	253	2	Q6J981	Q6J981 arabidopsis	635	10	1.8	760	2	Q86A14	Q86A14 dictyosteli
563	1.8	254	2	Q7SEA4	Q7SEA4 neurospora	636	10	1.8	761	2	Q7S5H9	Q7S5H9 dictyosteli
564	1.8	254	2	Q9W012	Q9W012 drosophila	637	10	1.8	788	2	Q9W180	Q9W180 drosophila
565	1.8	255	2	Q23112	Q23112 arabidopsis	638	10	1.8	791	1	CUDA_DICDI	Q00841 dictyosteli
566	1.8	260	2	Q73T85	Q73T85 mycobacteri	639	10	1.8	799	2	Q7YU4	Q7YU4 cryptospori
567	1.8	263	1	SGS3_DROYA	P13728 drosophila	640	10	1.8	805	2	Q8MLP2	Q8MLP2 drosophila
568	1.8	272	2	Q86EW5	Q86EW5 schistosoma	641	10	1.8	812	2	Q86AW3	Q86AW3 dictyosteli
569	1.8	274	2	Q86KZ5	Q86KZ5 dictyosteli	642	10	1.8	816	2	Q28331	Q28331 archaeoglob
570	1.8	275	2	Q8MNV5	Q8MNV5 dictyosteli	643	10	1.8	816	2	Q86HD8	Q86HD8 dictyosteli
571	1.8	294	2	Q7PK72	Q7PK72 anopheles g	644	10	1.8	817	2	Q7QAN0	Q7QAN0 anopheles g
572	1.8	307	1	SGS3_DROME	P02840 drosophila	645	10	1.8	820	2	Q754J6	Q754J6 ashbya gos
573	1.8	311	2	Q7O5Z0	Q7O5Z0 anopheles g	646	10	1.8	828	2	Q75JJO	Q75JJO dictyosteli
574	1.8	335	2	Q9ZTC1	Q9ZTC1 arabidopsis	647	10	1.8	835	2	Q75JY0	Q75JY0 dictyosteli
575	1.8	344	2	Q9N311	Q9N311 caenorhabdi	648	10	1.8	838	2	Q8R3S1	Q8R3S1 mus musculu
576	1.8	348	2	Q86JW4	Q86JW4 dictyosteli	649	10	1.8	853	2	Q7RWY1	Q7RWY1 neurospora
577	1.8	356	2	Q818P7	Q818P7 anopheles g	650	10	1.8	857	2	Q6C1I6	Q6C1I6 varrowia li
578	1.8	356	2	Q7PG96	Q7PG96 anopheles g	651	10	1.8	888	2	Q6KAT0	Q6KAT0 mus musculu
579	1.8	368	2	Q7PK74	Q7PK74 anopheles g	652	10	1.8	904	2	Q86HL7	Q86HL7 dictyosteli
580	1.8	391	2	Q7FAX5	Q7FAX5 anopheles g	653	10	1.8	908	2	Q80YR4	Q80YR4 mus musculu
581	1.8	393	2	Q13495	Q13495 neocallimas	654	10	1.8	911	2	Q86H89	Q86H89 dictyosteli
582	1.8	393	2	Q9ZVZ4	Q9ZVZ4 arabidopsis	655	10	1.8	919	2	Q91BB5	Q91BB5 spodoptera
583	1.8	395	2	Q9GNB6	Q9GNB6 arabidopsis	656	10	1.8	921	2	Q869X3	Q869X3 dictyosteli
584	1.8	403	2	Q80UL4	Q80UL4 mus musculu	657	10	1.8	923	2	Q17970	Q17970 caenorhabdi
585	1.8	411	2	Q9F8S2	Q9F8S2 oryza sativ	658	10	1.8	941	2	Q869Y0	Q869Y0 dictyosteli
586	1.8	413	2	Q7PK75	Q7PK75 anopheles g	659	10	1.8	947	2	Q15739	Q15739 dictyosteli
587	1.8	417	2	Q6V4B7	Q6V4B7 drosophila	660	10	1.8	970	2	Q86AL6	Q86AL6 dictyosteli
588	1.8	429	2	Q7ZUE5	Q7ZUE5 brachydanio	661	10	1.8	979	2	Q7KWS3	Q7KWS3 dictyosteli
589	1.8	431	2	Q22454	Q22454 caenorhabdi	662	10	1.8	990	2	Q75JK7	Q75JK7 dictyosteli
590	1.8	434	2	Q22458	Q22458 caenorhabdi	663	10	1.8	999	2	Q86UE5	Q86UE5 homo sapien
591	1.8	435	2	Q45851	Q45851 caenorhabdi	664	10	1.8	1008	2	Q74672	Q74672 pneumocysti
592	1.8	450	2	Q7S732	Q7S732 neurospora	665	10	1.8	1012	2	Q81I31	Q81I31 caenorhabdi
593	1.8	451	1	MYBH_DICDI	P34127 dictyosteli	666	10	1.8	1014	2	Q17505	Q17505 caenorhabdi
594	1.8	451	2	Q965P9	Q965P9 caenorhabdi	667	10	1.8	1017	2	Q74670	Q74670 pneumocysti
595	1.8	465	2	Q9BIP2	Q9BIP2 cooperia pu	668	10	1.8	1022	2	Q74671	Q74671 pneumocysti
596	1.8	468	2	Q7XMQ1	Q7XMQ1 oryza sativ	669	10	1.8	1026	2	Q74669	Q74669 pneumocysti
597	1.8	474	2	Q7JP78	Q7JP78 caenorhabdi	670	10	1.8	1028	2	Q74668	Q74668 pneumocysti
598	1.8	476	2	Q86J73	Q86J73 dictyosteli	671	10	1.8	1030	2	Q01829	Q01829 pneumocysti
599	1.8	480	2	Q7SHC6	Q7SHC6 neurospora	672	10	1.8	1050	2	Q9U6U8	Q9U6U8 dictyosteli
600	1.8	480	2	Q873J7	Q873J7 neurospora	673	10	1.8	1059	2	Q7S416	Q7S416 neurospora
601	1.8	489	2	Q8MTX9	Q8MTX9 drosophila	674	10	1.8	1064	2	Q8N3D8	Q8N3D8 homo sapien
602	1.8	489	2	Q7Y187	Q7Y187 oryza sativ	675	10	1.8	1080	2	Q6QR24	Q6QR24 trypanosoma
603	1.8	490	2	Q6V6N2	Q6V6N2 drosophila	676	10	1.8	1083	2	Q813G1	Q813G1 plasmodium
604	1.8	490	2	Q6V6N4	Q6V6N4 drosophila	677	10	1.8	1089	2	Q86KP4	Q86KP4 dictyosteli
605	1.8	490	2	Q6V6N6	Q6V6N6 drosophila	678	10	1.8	1109	2	Q8XMS8	Q8XMS8 clostridium
606	1.8	522	2	Q86JZ2	Q86JZ2 dictyosteli	679	10	1.8	1164	2	Q7YU46	Q7YU46 drosophila
607	1.8	525	2	Q8B3F1	Q8B3F1 mus musculu	680	10	1.8	1208	2	Q9NGW7	Q9NGW7 dictyosteli
608	1.8	547	2	Q81943	Q81943 drosophila	681	10	1.8	1219	2	Q95XG5	Q95XG5 caenorhabdi
609	1.8	549	2	Q81Q59	Q81Q59 drosophila	682	10	1.8	1278	2	Q7S2S7	Q7S2S7 neurospora
610	1.8	560	2	Q7P847	Q7P847 anopheles g	683	10	1.8	1307	2	Q9V4J6	Q9V4J6 drosophila
611	1.8	579	2	Q8MT56	Q8MT56 drosophila	684	10	1.8	1325	2	Q75JA1	Q75JA1 dictyosteli
612	1.8	579	2	Q9VB46	Q9VB46 drosophila	685	10	1.8	1348	2	Q86KE2	Q86KE2 dictyosteli
613	1.8	581	2	Q81925	Q81925 hyphantria	686	10	1.8	1377	2	O42695	O42695 candida alb
614	1.8	582	2	Q24160	Q24160 drosophila	687	10	1.8	1390	2	O77033	O77033 dictyosteli
615	1.8	626	2	Q9NDD1	Q9NDD1 leishmania	688	10	1.8	1418	2	Q86HT4	Q86HT4 dictyosteli

689	10	1.8	1457	2	Q81S20	Q81a20 dictyosteli	762	9	1.6	241	2	Q9VDR0	Q9vdr0 drosophila
690	10	1.8	1478	2	Q75J11	Q75j11 dictyosteli	763	9	1.6	245	2	Q872C8	Q872c8 neurospora
691	10	1.8	1541	2	Q9W8M7	Q9w8m7 drosophila	764	9	1.6	248	2	Q9BIQ6	Q9biq6 cooperia pu
692	10	1.8	1543	2	Q9GV71	Q9gv71 dictyosteli	765	9	1.6	248	2	Q9BIQ8	Q9biq8 cooperia pu
693	10	1.8	1553	2	Q95TR0	Q95tr0 drosophila	766	9	1.6	251	2	Q6CHT1	Q6cht1 yarrowia li
694	10	1.8	1570	1	P3K1_DICD1	P54673 dictyosteli	767	9	1.6	252	2	Q8BQ75	Q8bq75 mus musculus
695	10	1.8	1616	2	Q7KU53	Q7kub3 drosophila	768	9	1.6	257	2	Q8BQ75	Q8bq75 mus musculus
696	10	1.8	1616	2	Q9V5J0	Q9v5j0 drosophila	769	9	1.6	261	2	Q6IH30	Q6ih30 drosophila
697	10	1.8	1629	2	Q75J33	Q75j33 dictyosteli	770	9	1.6	273	2	Q65XK9	Q65xk9 oryza sativ
698	10	1.8	1716	2	Q917Q8	Q917q8 drosophila	771	9	1.6	278	2	Q6FIQ7	Q6fiq7 candida gla
699	10	1.8	1726	2	Q6XHB2	Q6xbh2 dictyosteli	772	9	1.6	278	2	Q7Q5F5	Q7q5f5 anopheles g
700	10	1.8	1740	2	Q9HCB0	Q9hcj0 homo sapien	773	9	1.6	279	2	Q7Q5F5	Q7q5f5 anopheles g
701	10	1.8	1761	2	Q6ZQH8	Q6zqh8 mus musculus	774	9	1.6	287	2	Q7Q5F5	Q7q5f5 anopheles g
702	10	1.8	1845	2	Q75I33	Q75i33 neurospora	775	9	1.6	287	2	Q7Q5F5	Q7q5f5 anopheles g
703	10	1.8	1880	2	Q8MP27	Q8mp27 dictyosteli	776	9	1.6	290	2	Q9Y1H4	Q9y1h4 dictyosteli
704	10	1.8	1887	2	Q8SSY6	Q8ssy6 dictyosteli	777	9	1.6	294	2	Q8C6E5	Q8c6e5 mus musculus
705	10	1.8	1912	2	Q9SVS9	Q9svs9 drosophila	778	9	1.6	308	2	Q9CSZ3	Q9csz3 mus musculus
706	10	1.8	1969	2	Q7SEZ5	Q7se25 neurospora	779	9	1.6	309	2	Q7SOP9	Q7sop9 neurospora
707	10	1.8	2062	2	Q9SFP6	Q9spn6 dictyosteli	780	9	1.6	315	2	Q81QQ1	Q81qq1 drosophila
708	10	1.8	2112	2	Q61DH3	Q61dh3 drosophila	781	9	1.6	316	2	Q6FXF0	Q6fxf0 candida gla
709	10	1.8	2140	2	Q7SG46	Q7sg46 neurospora	782	9	1.6	316	2	Q6ZJ42	Q6zj42 oryza sativ
710	10	1.8	2181	2	Q9VRA6	Q9vra6 drosophila	783	9	1.6	318	2	Q9N4M7	Q9n4m7 caenorhabdi
711	10	1.8	2507	2	Q81TH8	Q81hk8 dictyosteli	784	9	1.6	321	2	Q17977	Q17977 caenorhabdi
712	10	1.8	2646	2	Q6XHA6	Q6xha6 dictyosteli	785	9	1.6	327	2	Q717U3	Q717u3 nicotiana t
713	10	1.8	2800	2	Q6XHB1	Q6xhb1 dictyosteli	786	9	1.6	327	2	Q9SXX1	Q9sxx1 nicotiana t
714	10	1.8	2833	2	Q9VP13	Q9vp13 drosophila	787	9	1.6	327	2	Q8BY98	Q8by98 m mus muscu
715	10	1.8	2929	2	Q86JG5	Q86jg5 dictyosteli	788	9	1.6	332	1	P111_MOUSE	Q9c2x5 mus musculus
716	10	1.8	3135	2	Q7KHU2	Q7kuh2 drosophila	789	9	1.6	354	2	Q86HY8	Q86hy8 dictyosteli
717	10	1.8	3166	2	Q9W3Z0	Q9w3z0 drosophila	790	9	1.6	359	2	Q90351	Q90351 cornuix co
718	10	1.8	3455	2	Q6R5A9	Q6r5a9 tenebrio mo	791	9	1.6	364	2	Q9YZ52	Q9yz52 chimpanzee
719	10	1.8	4498	2	Q9W2Z3	Q9w2z3 drosophila	792	9	1.6	368	2	Q9YZ45	Q9yz45 chimpanzee
720	10	1.8	8905	2	Q7KR73	Q7kr73 drosophila	793	9	1.6	389	2	Q9Y6R9	Q9y6r9 drosophila
721	9	1.6	41	2	Q7R9M4	Q7r9m4 plasmodium	794	9	1.6	390	2	Q6CEI2	Q6cei2 yarrowia li
722	9	1.6	48	2	Q87172	Q87172 chimpanzee	795	9	1.6	390	2	Q81927	Q81927 hyphantria
723	9	1.6	77	2	Q86NZ8	Q86nz8 drosophila	796	9	1.6	396	2	Q8H880	Q8h880 oryza sativ
724	9	1.6	93	2	Q9VUA6	Q9vua6 drosophila	797	9	1.6	408	2	Q6W4E5	Q6w4e5 drosophila
725	9	1.6	105	2	Q7S3N9	Q7s3n9 neurospora	798	9	1.6	408	2	Q6W4E7	Q6w4e7 drosophila
726	9	1.6	106	2	Q9NIQ6	Q9niq6 trypanosoma	799	9	1.6	408	2	Q6W4E9	Q6w4e9 drosophila
727	9	1.6	109	2	Q9VTF2	Q9vtf2 drosophila	800	9	1.6	408	2	Q6W4F2	Q6w4f2 drosophila
728	9	1.6	117	2	Q9BJP9	Q9bjp9 plasmodium	801	9	1.6	408	2	Q6W4G1	Q6w4g1 drosophila
729	9	1.6	120	2	Q04420	Q04420 chimpanzee	802	9	1.6	408	2	Q6W4G4	Q6w4g4 drosophila
730	9	1.6	122	2	Q81PE4	Q81pe4 drosophila	803	9	1.6	408	2	Q6W4G7	Q6w4g7 drosophila
731	9	1.6	124	2	Q7RUG3	Q7rug3 neurospora	804	9	1.6	408	2	Q6W4H2	Q6w4h2 drosophila
732	9	1.6	127	2	Q61035	Q61035 trypanosoma	805	9	1.6	409	2	Q8MKM3	Q8mkm3 drosophila
733	9	1.6	131	2	Q61J65	Q61j65 drosophila	806	9	1.6	409	2	Q86JU3	Q86ju3 dictyosteli
734	9	1.6	135	2	Q01623	Q01623 pneumocysti	807	9	1.6	410	2	Q6Y144	Q6y144 lactuca sal
735	9	1.6	142	2	Q26943	Q26943 trypanosoma	808	9	1.6	413	2	Q86I25	Q86i25 dictyosteli
736	9	1.6	145	2	Q6RUB0	Q6rub0 chimpanzee	809	9	1.6	413	2	Q8Y126	Q8y126 lactuca sat
737	9	1.6	147	2	Q6RUD2	Q6rud2 chimpanzee	810	9	1.6	422	2	Q86AN7	Q86an7 dictyosteli
738	9	1.6	147	2	Q6RUD7	Q6rud7 chimpanzee	811	9	1.6	426	2	Q6Y159	Q6y159 lactuca sat
739	9	1.6	147	2	Q6RUE4	Q6rue4 chimpanzee	812	9	1.6	433	1	RTC1_DICD1	Q15746 dictyosteli
740	9	1.6	158	2	Q68KL8	Q68kl8 nasonia vit	813	9	1.6	435	2	Q9FPA2	Q9fpa2 oryza sativ
741	9	1.6	166	2	Q73V43	Q73v43 mycobacteri	814	9	1.6	438	2	Q8LH48	Q8lh48 oryza sativ
742	9	1.6	168	2	Q61IG4	Q61ig4 drosophila	815	9	1.6	440	2	Q7S0A4	Q7s0a4 neurospora
743	9	1.6	170	2	Q8TIF7	Q8tif7 dictyosteli	816	9	1.6	441	2	Q8T1W3	Q8t1w3 dictyosteli
744	9	1.6	172	2	Q96QT0	Q96qt0 homo sapien	817	9	1.6	444	2	Q6MVL1	Q6mvl1 neurospora
745	9	1.6	178	2	Q922K0	Q922k0 mus musculus	818	9	1.6	453	2	Q12945	Q12945 gallus gall
746	9	1.6	184	2	Q6K966	Q6k966 oryza sativ	819	9	1.6	461	2	Q9FUL9	Q9ful9 zea mays (m
747	9	1.6	200	2	Q9N2S0	Q9n2s0 trypanosoma	820	9	1.6	464	2	Q871Q8	Q871q8 neurospora
748	9	1.6	202	2	Q9AZX9	Q9azx9 bacterioph	821	9	1.6	475	2	Q8MSD6	Q8msd6 drosophila
749	9	1.6	202	2	Q9CGQ0	Q9cgq0 lactococcus	822	9	1.6	479	2	Q870W9	Q870w9 neurospora
750	9	1.6	207	2	Q61045	Q61045 trypanosoma	823	9	1.6	483	2	Q8T8X4	Q8t8x4 drosophila
751	9	1.6	208	2	Q9U0A0	Q9u0a0 plasmodium	824	9	1.6	483	2	Q9VK10	Q9vk10 drosophila
752	9	1.6	209	2	Q6A279	Q6a279 haemophilus	825	9	1.6	484	2	Q7P2E5	Q7p2e5 anopheles g
753	9	1.6	216	2	Q6X857	Q6x857 spodoptera	826	9	1.6	484	2	Q87G85	Q87ges anopheles g
754	9	1.6	220	2	Q7S6T5	Q7s6t5 neurospora	827	9	1.6	488	1	UZ1P_DRONE	P10379 drosophila
755	9	1.6	224	2	Q8LGR8	Q8lgr8 lycopersico	828	9	1.6	490	2	Q6I5U8	Q6i5u8 oryza sativ
756	9	1.6	224	2	Q8VZ00	Q8vz00 arabidopsis	829	9	1.6	492	1	MEC2_RAT	Q00566 rattus norv
757	9	1.6	224	2	Q91Q83	Q91q83 arabidopsis	830	9	1.6	492	2	Q6PBT9	Q6pbt9 brachydanio
758	9	1.6	232	2	Q8NI48	Q8ni48 homo sapien	831	9	1.6	494	2	Q9VWU0	Q9vwu0 drosophila
759	9	1.6	232	2	Q27423	Q27423 drosophila	832	9	1.6	496	2	Q7S6G4	Q7s6g4 neurospora
760	9	1.6	233	2	Q9U0C2	Q9u0c2 plasmodium	833	9	1.6	497	2	Q7SHV7	Q7shv7 neurospora
761	9	1.6	237	2	Q81RY4	Q81ry4 drosophila	834	9	1.6	499	2	Q95QK9	Q95qk9 caenorhabdi

835	9	1.6	505	2	Q7SHM1	Q7ehm1 neurospora	908	9	1.6	306	1	YPG1	DICSP	P10511 dictyosteli
836	9	1.6	509	2	Q94988	Q94888 drosophila	909	9	1.6	308	2	Q7S9B2		Q7S9b2 neurospora
837	9	1.6	509	2	Q7KM08	Q7km08 drosophila	910	9	1.6	914	2	Q86K47		Q86k47 dictyosteli
838	9	1.6	511	2	Q9VZH5	Q9vzh5 drosophila	911	9	1.6	919	2	Q8QS30		Q8q30 pongine her
839	9	1.6	514	2	Q7S7J2	Q7s7j2 neurospora	912	9	1.6	920	2	Q86KS1		Q86ks1 dictyosteli
840	9	1.6	519	2	Q7YTR7	Q7ytr7 caenorhabdi	913	9	1.6	929	2	Q8SSQ3		Q8ssq3 dictyosteli
841	9	1.6	522	2	Q9XTK9	Q9xtk9 drosophila	914	9	1.6	941	2	Q9VXA2		Q9vxa2 drosophila
842	9	1.6	523	2	Q90S23	Q90s23 chimpanzee	915	9	1.6	943	1	RNGB	DICDI	Q7m3s9 dictyosteli
843	9	1.6	525	2	Q7SBB9	Q7sbb9 neurospora	916	9	1.6	950	2	Q8MQN5		Q8mqn5 drosophila
844	9	1.6	526	2	Q6NVT9	Q6nv99 homo sapien	917	9	1.6	968	2	Q8IR41		Q8ir41 drosophila
845	9	1.6	526	2	Q9UK58	Q9uk58 homo sapien	918	9	1.6	980	2	Q75JK3		Q75jk3 dictyosteli
846	9	1.6	527	2	Q9V9J6	Q9v9j6 drosophila	919	9	1.6	989	1	PTP3	DICDI	P54637 dictyosteli
847	9	1.6	527	2	Q9R1Q2	Q9r1q2 rattus norv	920	9	1.6	1004	2	Q8MP26		Q8mp26 dictyosteli
848	9	1.6	532	2	Q8R5H9	Q8r5h9 mus musculu	921	9	1.6	1014	2	Q8SSY2		Q8ssy2 dictyosteli
849	9	1.6	534	2	Q7Q3D3	Q7q3d3 anopheles g	922	9	1.6	1024	2	Q75JA0		Q75ja0 dictyosteli
850	9	1.6	535	2	Q95QK8	Q95qk8 caenorhabdi	923	9	1.6	1043	2	Q8SSW7		Q8ssw7 dictyosteli
851	9	1.6	537	2	Q8IFC6	Q8if66 drosophila	924	9	1.6	1055	2	Q75J96		Q75j96 dictyosteli
852	9	1.6	540	2	Q7PS94	Q7ps94 anopheles g	925	9	1.6	1093	2	Q6NCH8		Q6nch8 rhodopeudo
853	9	1.6	545	2	Q8T1B3	Q8t1b3 dictyosteli	926	9	1.6	1135	2	Q8SSU8		Q8ssu8 dictyosteli
854	9	1.6	546	2	Q8SXG6	Q8sxg6 drosophila	927	9	1.6	1144	2	Q86JE2		Q86je2 dictyosteli
855	9	1.6	556	2	Q9VFW1	Q9vfw1 drosophila	928	9	1.6	1147	2	Q95PH8		Q95ph8 dictyosteli
856	9	1.6	567	1	CHI3 CANAL	P40954 candida alb	929	9	1.6	1148	2	Q7QGP4		Q7Qgp4 dictyosteli
857	9	1.6	578	2	Q8BPP4	Q8bpp4 mus musculu	930	9	1.6	1163	2	Q8GAR1		Q8gar1 dictyosteli
858	9	1.6	583	2	Q6K5Q2	Q6k5q2 oryza sativ	931	9	1.6	1175	2	Q9VRL7		Q9vrl7 drosophila
859	9	1.6	584	2	Q8G9Z3	Q8g9z3 dictyosteli	932	9	1.6	1176	2	Q8G9T7		Q8g9t7 dictyosteli
860	9	1.6	584	2	Q8K480	Q8k480 mus musculu	933	9	1.6	1205	2	Q86JG9		Q86jg9 dictyosteli
861	9	1.6	587	2	Q75JZ1	Q75jz1 dictyosteli	934	9	1.6	1222	2	Q7PPC0		Q7ppc0 anopheles g
862	9	1.6	607	2	Q9VXZ2	Q9vxz2 drosophila	935	9	1.6	1223	2	Q7KWX7		Q7kwx7 dictyosteli
863	9	1.6	610	2	Q6Y0X6	Q6y0x6 mus musculu	936	9	1.6	1241	2	Q75JCO		Q75jco dictyosteli
864	9	1.6	615	2	Q84SM4	Q84sm4 oryza sativ	937	9	1.6	1259	2	Q8G9X4		Q8g9x4 dictyosteli
865	9	1.6	617	1	ESR1 ICTPU	Q9yhz7 ictalurus p	938	9	1.6	1324	2	Q8MMQ2		Q8mmq2 dictyosteli
866	9	1.6	619	2	Q76P26	Q76p26 dictyosteli	939	9	1.6	1329	2	Q75JPO		Q75jpo dictyosteli
867	9	1.6	622	2	Q7S9H1	Q7s9h1 neurospora	940	9	1.6	1369	1	NFAS	CHICK	Q42414 gallus gall
868	9	1.6	622	2	Q9P389	Q9p389 neurospora	941	9	1.6	1407	2	Q9VB65		Q9vb65 drosophila
869	9	1.6	622	2	Q6Y141	Q6y141 lactuca sat	942	9	1.6	1408	1	SERR	DROME	P18168 drosophila
870	9	1.6	624	2	Q19780	Q19780 caenorhabdi	943	9	1.6	1458	2	Q7R2A6		Q7r2a6 neurospora
871	9	1.6	638	1	UBQ2 MOUSE	Q9qzm0 mus musculu	944	9	1.6	1464	2	Q61802		Q61802 caenorhabdi
872	9	1.6	639	2	Q8G5X5	Q8g5x5 dictyosteli	945	9	1.6	1470	2	Q9VYK5		Q9vyk5 drosophila
873	9	1.6	653	2	Q7S3H6	Q7s3h6 neurospora	946	9	1.6	1495	2	Q8GAU5		Q8gau5 dictyosteli
874	9	1.6	655	2	Q7KKR0	Q7kkr0 drosophila	947	9	1.6	1518	2	Q21442		Q21442 caenorhabdi
875	9	1.6	658	2	Q7PNH9	Q7pnh9 anopheles g	948	9	1.6	1542	2	Q9VE07		Q9ve07 drosophila
876	9	1.6	662	2	Q8MQZ9	Q8mqz9 drosophila	949	9	1.6	1557	2	Q76NT8		Q76nt8 dictyosteli
877	9	1.6	662	2	Q9VGD0	Q9vgd0 drosophila	950	9	1.6	1752	2	Q9AE52		Q9ae52 ruminococcu
878	9	1.6	665	2	Q86HZ3	Q86hz3 dictyosteli	951	9	1.6	1789	2	Q8T145		Q8t145 dictyosteli
879	9	1.6	670	2	Q6PW12	Q6pw12 gallus gall	952	9	1.6	1795	2	Q76894		Q76894 drosophila
880	9	1.6	679	2	Q69HN9	Q69hn9 ciona intes	953	9	1.6	1806	2	Q8G9R4		Q8g9r4 dictyosteli
881	9	1.6	682	2	Q6DJD2	Q6ddj2 xenopus lae	954	9	1.6	1837	2	Q8IKF1		Q8ikf1 plasmodium
882	9	1.6	704	2	Q86IW5	Q86iw5 dictyosteli	955	9	1.6	1837	2	Q8G9Q5		Q8g9q5 dictyosteli
883	9	1.6	708	2	Q22806	Q22806 caenorhabdi	956	9	1.6	1847	2	Q7R5Z0		Q7r5z0 giardia lam
884	9	1.6	709	2	Q86IX4	Q86ix4 dictyosteli	957	9	1.6	1997	2	Q8ISL6		Q8isl6 plasmodium
885	9	1.6	714	2	Q8T210	Q8t210 dictyosteli	958	9	1.6	2087	2	Q8MXL2		Q8mxl2 leishmania
886	9	1.6	714	2	Q8T210	Q8t210 dictyosteli	959	9	1.6	2335	2	Q7VYR5		Q7vyr5 cryptospori
887	9	1.6	740	2	Q8FLY9	Q8fly9 corynebacte	960	9	1.6	2602	2	Q9VZ45		Q9vz45 drosophila
888	9	1.6	741	2	Q86HM4	Q86hm4 dictyosteli	961	9	1.6	2691	2	Q8ILS2		Q8ils2 plasmodium
889	9	1.6	745	2	Q86CR5	Q86cr5 dictyosteli	962	9	1.6	2752	2	Q6LFI9		Q6lfi9 plasmodium
890	9	1.6	745	2	Q86KJ5	Q86kj5 dictyosteli	963	9	1.6	3633	2	Q8IHL0		Q8ihl0 dictyosteli
891	9	1.6	759	2	Q05143	Q05143 ruminococcu	964	9	1.6	3672	1	LML2	CAEEL	Q21313 caenorhabdi
892	9	1.6	760	2	Q6B516	Q6b516 dictyosteli	965	9	1.6	3704	2	P91904		P91904 caenorhabdi
893	9	1.6	762	2	Q8BAG3	Q8bag3 dictyosteli	966	9	1.6	4118	2	Q8GTF2		Q8gtf2 dictyosteli
894	9	1.6	767	2	Q86KE5	Q86ke5 dictyosteli	967	9	1.6	4138	2	Q8L1Y3		Q8l1y3 plasmodium
895	9	1.6	777	2	Q7K7G5	Q7k7g5 drosophila	968	9	1.6	7524	2	Q6PEZ0		Q6pezo mus musculu
896	9	1.6	780	1	A4_TETPL	Q73683 tetraodon f	969	9	1.6	9234	2	Q7KTP5		Q7ktp5 drosophila
897	9	1.6	786	2	Q86L03	Q86l03 dictyosteli	970	8	1.4	32	2	Q9S442		Q9s442 neisseria m
898	9	1.6	787	2	Q817P3	Q817p3 dictyosteli	971	8	1.4	47	2	Q8J4H6		Q8j4h6 chimpanzee
899	9	1.6	794	2	Q7KU09	Q7ku09 drosophila	972	8	1.4	47	2	Q87205		Q87205 chimpanzee
900	9	1.6	798	2	Q7S386	Q7s386 dictyosteli	973	8	1.4	47	2	Q87209		Q87209 chimpanzee
901	9	1.6	799	2	Q8GA08	Q8ga08 dictyosteli	974	8	1.4	47	2	Q87260		Q87260 chimpanzee
902	9	1.6	801	2	Q9W3K6	Q9w3k6 drosophila	975	8	1.4	47	2	Q87261		Q87261 chimpanzee
903	9	1.6	811	2	Q86I20	Q86i20 dictyosteli	976	8	1.4	48	2	Q7KJB8		Q7kjb8 plasmodium
904	9	1.6	814	2	Q80UR5	Q80ur5 mus musculu	977	8	1.4	48	2	Q9N6A3		Q9n6a3 plasmodium
905	9	1.6	867	2	Q8S2E7	Q8sze7 drosophila	978	8	1.4	48	2	Q8J4I4		Q8j4i4 chimpanzee
906	9	1.6	876	2	Q75JU2	Q75ju2 dictyosteli	979	8	1.4	48	2	Q87164		Q87164 chimpanzee
907	9	1.6	879	2	Q76874	Q76874 drosophila	980	8	1.4	48	2	Q87176		Q87176 chimpanzee



881 8 1.4 48 2 Q87193 chimpanzee  
 882 8 1.4 48 2 Q87194 chimpanzee  
 883 8 1.4 48 2 Q87201 chimpanzee  
 884 8 1.4 48 2 Q87222 chimpanzee  
 885 8 1.4 48 2 Q87255 chimpanzee  
 886 8 1.4 48 2 Q87256 chimpanzee  
 887 8 1.4 48 2 Q87257 chimpanzee  
 888 8 1.4 48 2 Q87258 chimpanzee  
 889 8 1.4 48 2 Q87265 chimpanzee  
 890 8 1.4 48 2 Q87266 chimpanzee  
 891 8 1.4 48 2 Q87269 chimpanzee  
 892 8 1.4 48 2 Q87274 chimpanzee  
 993 8 1.4 50 2 Q9N120 plasmodium  
 994 8 1.4 51 2 Q87187 chimpanzee  
 995 8 1.4 51 2 Q87191 chimpanzee  
 996 8 1.4 51 2 Q87202 chimpanzee  
 997 8 1.4 51 2 Q87203 chimpanzee  
 998 8 1.4 51 2 Q87213 chimpanzee  
 999 8 1.4 51 2 Q87263 chimpanzee  
 1000 8 1.4 51 2 Q87270 chimpanzee

## ALIGNMENTS

RESULT 1  
 Q7DB77  
 ID Q7DB77 PRELIMINARY; PRT; 558 AA.  
 AC Q7DB77 Q7A9Q1;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Putative translocated intimin receptor protein (Translocated intimin receptor Tir).  
 DE Name=tir; OrderedLocusNames=EC84561, Z5112;  
 GN Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=83334;  
 RN [1]\_TaxID=83334;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;  
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;  
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7;  
 RL Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T., Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL; AE005595; AAG58825.1; -;  
 DR EMBL; AF002566; BAB37984.1; -;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR003536; Tir receptor.  
 DR Pfam; PF07489; Tir receptor\_C; 1.  
 DR Pfam; PF03549; Tir receptor\_M; 1.  
 DR Pfam; PF07490; Tir receptor\_N; 1.  
 DR PRINTS; PR01370; TRENINIMINR.  
 SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4B4AA1 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 558;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MPINLGNPNVNSIIPAPPLPSQTDCAGRGOLINSTGLGSRALFTPVNSMADSGD 60  
 Db 1 MPINLGNPNVNSIIPAPPLPSQTDCAGRGOLINSTGLGSRALFTPVNSMADSGD 60  
 Qy 61 NRASDVPLGNPMLAAASEITLNDGFEVLHDHGLDPLNRIQIGSSVFRVETQEDGKHIA 120  
 Db 61 NRASDVPLGNPMLAAASEITLNDGFEVLHDHGLDPLNRIQIGSSVFRVETQEDGKHIA 120  
 Qy 121 VQONGVETSVLSDQEVARYLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180  
 Db 121 VQONGVETSVLSDQEVARYLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180  
 Qy 181 ELLPEKGTGSGKAGESKVGELRESNSGAENTTQTSTSLRSPDKLWALGTAT 240  
 Db 181 ELLPEKGTGSGKAGESKVGELRESNSGAENTTQTSTSLRSPDKLWALGTAT 240  
 Qy 241 GLIGLAATGIYOALALTPEDSPPTTDPDAASATETATRDQLTKEAFQNPDKVNIDE 300  
 Db 241 GLIGLAATGIYOALALTPEDSPPTTDPDAASATETATRDQLTKEAFQNPDKVNIDE 300  
 Qy 301 LGNAIPSGVLKDDVVANIEQAKAAGEAKQQAENNAQAOKKYDEQAQKEELKVSSG 360  
 Db 301 LGNAIPSGVLKDDVVANIEQAKAAGEAKQQAENNAQAOKKYDEQAQKEELKVSSG 360  
 Qy 361 AGYGLSGALIILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTSARTVENKPNNTPAQG 420  
 Db 361 AGYGLSGALIILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTSARTVENKPNNTPAQG 420  
 Qy 421 NVDTPGSEDWTESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480  
 Db 421 NVDTPGSEDWTESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480  
 Qy 481 NMGNTDSVYVSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLTGGSN 540  
 Db 481 NMGNTDSVYVSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLTGGSN 540  
 Qy 541 SAVNTSNNPPAGSHRFV 558  
 Db 541 SAVNTSNNPPAGSHRFV 558  
 RESULT 2  
 Q9R396 PRELIMINARY; PRT; 558 AA.  
 ID Q9R396  
 AC Q9R396;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Translocated intimin receptor Tir (L0027).  
 GN Name=tir;  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]\_TaxID=562;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=86/24;  
 RX MEDLINE=99242825; PubMed=10225900;  
 RA DeVinney R., Stein M., Reinscheid D., Abe A., Ruschkowski S., Finlay B.B.;  
 RA "Enterohaemorrhagic Escherichia coli O157:H7 produces Tir, which is translocated to the host cell membrane but is not tyrosine phosphorylated.";  
 RL Infect. Immun. 67:2389-2398(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC43895; and EDL933;  
 RX MEDLINE=98339885; PubMed=9673266;  
 RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,

RA Kaper J.B., Blattner F.R.;  
 RT "Molecular evolution of a pathogenicity island from enterohemorrhagic  
 RT Escherichia coli O157:H7", (1998).  
 RL Infect. Immun. 66:3810-3817 (1998).  
 DR EMBL; AF125993; AAC29391.1; -.  
 DR EMBL; AF071034; AAC31506.1; -.  
 DR PIR; A98199; A98199.  
 DR PIR; A98199; A98199.  
 DR HSP; O9KWH9; 1F02.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR003536; Tir receptor.  
 DR Pfam; PF07489; Tir\_receptor\_C; 1.  
 DR Pfam; PF03549; Tir\_receptor\_M; 1.  
 DR Pfam; PF07490; Tir\_receptor\_N; 1.  
 DR PRINTS; PR01370; TRNSINTMINR.  
 KW Receptor.  
 SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4B4AA1 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 558;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNNSIPPAPLPSTQDAGGGRGQLINSTGPGSRALFTPVNSMADSGD 60  
 DB 1 MPIGNLGNPNVNNNSIPPAPLPSTQDAGGGRGQLINSTGPGSRALFTPVNSMADSGD 60  
 QY 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDGPDLTNLRQIGSSVFRVETQEDGKHA 120  
 DB 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDGPDLTNLRQIGSSVFRVETQEDGKHA 120  
 QY 121 VQQRNGVETSVLSDOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQIL 180  
 DB 121 VQQRNGVETSVLSDOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQIL 180  
 QY 181 ELLEPKGTGSGKAGESKGVGELRESNCAENTTETQTSSTSSLRSDPKMLALGTAT 240  
 DB 181 ELLEPKGTGSGKAGESKGVGELRESNCAENTTETQTSSTSSLRSDPKMLALGTAT 240  
 QY 241 GLIGLAATGIVQALALTPEPDSPTTTPDPAASAATETATRDLTKEAFQNPQKNIDE 300  
 DB 241 GLIGLAATGIVQALALTPEPDSPTTTPDPAASAATETATRDLTKEAFQNPQKNIDE 300  
 QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQQAQKEELKVSSG 360  
 DB 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQQAQKEELKVSSG 360  
 QY 361 AGVLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVENKPNNTPAQG 420  
 DB 361 AGVLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVENKPNNTPAQG 420  
 QY 421 NVDTPGSEDTMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480  
 DB 421 NVDTPGSEDTMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480  
 QY 481 NMGNTDSVYTIQHPRTDTNGARLLGNPSAGIOSTYARLALSGGLRHDGGLTGGSN 540  
 DB 481 NMGNTDSVYTIQHPRTDTNGARLLGNPSAGIOSTYARLALSGGLRHDGGLTGGSN 540  
 QY 541 SAVNTSNPPAPGSHRFV 558  
 DB 541 SAVNTSNPPAPGSHRFV 558

## RESULT 3

O85506 PRELIMINARY; PRT; 558 AA.  
 AC O85506  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Translocated intimin receptor Tir.  
 GN Name:tir;

## RESULT 4

O7BHL5 PRELIMINARY; PRT; 547 AA.  
 AC O7BHL5  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Translocated intimin receptor Tir.  
 OC Citrobacter rodentium.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 RN NCBI\_TaxID=562;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=95SP2;  
 RX MEDLINE=99003184; PubMed=9784578;  
 RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;  
 RT "Translocated intimin receptors (Tir) of Shiga-toxicogenic Escherichia  
 RT coli isolates belonging to serogroups O26, O111, and O157 react with  
 RT sera from patients with hemolytic-uremic syndrome and exhibit marked  
 RT sequence heterogeneity";  
 RL Infect. Immun. 66:5580-5586 (1998).  
 DR EMBL; AF070067; AAC69314.1; -.  
 DR HSP; O9KWH9; 1F02.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR003536; Tir receptor.  
 DR Pfam; PF07489; Tir\_receptor\_C; 1.  
 DR Pfam; PF03549; Tir\_receptor\_M; 1.  
 DR Pfam; PF07490; Tir\_receptor\_N; 1.  
 DR PRINTS; PR01370; TRNSINTMINR.  
 KW Receptor.  
 SQ SEQUENCE 558 AA; 58175 MW; CA2CDDAC94527C2B CRC64;

Query Match 60.9%; Score 340; DB 2; Length 558;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNNSIPPAPLPSTQDAGGGRGQLINSTGPGSRALFTPVNSMADSGD 60  
 DB 1 MPIGNLGNPNVNNNSIPPAPLPSTQDAGGGRGQLINSTGPGSRALFTPVNSMADSGD 60  
 QY 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDGPDLTNLRQIGSSVFRVETQEDGKHA 120  
 DB 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDGPDLTNLRQIGSSVFRVETQEDGKHA 120  
 QY 121 VQQRNGVETSVLSDOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQIL 180  
 DB 121 VQQRNGVETSVLSDOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQIL 180  
 QY 181 ELLEPKGTGSGKAGESKGVGELRESNCAENTTETQTSSTSSLRSDPKMLALGTAT 240  
 DB 181 ELLEPKGTGSGKAGESKGVGELRESNCAENTTETQTSSTSSLRSDPKMLALGTAT 240  
 QY 241 GLIGLAATGIVQALALTPEPDSPTTTPDPAASAATETATRDLTKEAFQNPQKNIDE 300  
 DB 241 GLIGLAATGIVQALALTPEPDSPTTTPDPAASAATETATRDLTKEAFQNPQKNIDE 300  
 QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQQAQKEELKVSSG 360  
 DB 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQQAQKEELKVSSG 360  
 QY 361 AGVLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVENKPNNTPAQG 420  
 DB 361 AGVLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVENKPNNTPAQG 420  
 QY 421 NVDTPGSEDTMESRRSSMAST 441  
 DB 421 NVDTPGSEDTMESRRSSMAST 441

```
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=67825;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DS100;
RX MEDLINE=21437640; PubMed=11553577;
RX DOI=10.1128/IAI.69.10.6323-6335.2001;
RA Dong W., Li Y., Vallance B.A., Finlay B.B.;
RT "Locus of enterocyte effacement from Citrobacter rodentium: sequence
RT analysis and evidence for horizontal transfer among attaching and
RT effacing pathogens.";
RL Infect. Immun. 69:6323-6335 (2001).
DR EMBL; AF311901; AAL06376.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_receptor.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMINR.
KW Receptor.
SQ SEQUENCE 547 AA; 56283 MW; 02CAC6D625FA6EE1 CRC64;

Query Match 5.0%; Score 28; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.6e-18;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DQLTKEAFQNPQNDQKVNDELGNALPSG 308
DB 282 DQLTKEAFQNPQNDQKVNDELGNALPSG 309

RESULT 5
Q9ETI1
ID Q9ETI1 PRELIMINARY; PRT; 547 AA.
AC Q9ETI1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Translocated intimin receptor Tir.
OS Citrobacter rodentium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=67825;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1843-73T, and DS100;
RX MEDLINE=20553330; PubMed=1101562;
RA Luperchio S.A., Newman J.V., Dangier C.A., Schrenzel M.D.,
RA Brenner D.J., Steigerwalt A.G., Schauer D.B.;
RT "Citrobacter rodentium, the causative agent of transmissible murine
RT colonic hyperplasia, exhibits clonality; synonymy of C. rodentium and
RT mouse-pathogenic Escherichia coli.";
RL J. Clin. Microbiol. 38:4343-4350 (2000).
DR EMBL; AF301618; AAG40758.1; -.
DR EMBL; AF301617; AAG25642.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_receptor.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMINR.
KW Receptor.
SQ SEQUENCE 547 AA; 56270 MW; CB8318B301049C37 CRC64;

Query Match 5.0%; Score 28; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.6e-18;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DQLTKEAFQNPQNDQKVNDELGNALPSG 308
DB 282 DQLTKEAFQNPQNDQKVNDELGNALPSG 309

us-09-189-415b-11.oligo.rup

RESULT 6
Q9WXK1
ID Q9WXK1 PRELIMINARY; PRT; 547 AA.
AC Q9WXK1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MPEC;
RA Okutani A., Itoh K., Sasakawa C.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB026719; BAA77400.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_receptor.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMINR.
KW Receptor.
SQ SEQUENCE 547 AA; 56283 MW; 02CAC6D625FA6EE1 CRC64;

Query Match 5.0%; Score 28; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.6e-18;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DQLTKEAFQNPQNDQKVNDELGNALPSG 308
DB 282 DQLTKEAFQNPQNDQKVNDELGNALPSG 309

RESULT 7
O85508
ID O85508 PRELIMINARY; PRT; 538 AA.
AC O85508;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Translocated intimin receptor Tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EPSC87A;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (Tir) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586 (1998).
DR EMBL; AF070069; AAC69318.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_receptor.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
```

```
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor. 538 AA; 55602 MW; 447052A0E3214D6D CRC64;
SQ SEQUENCE 538 AA; 55602 MW; 447052A0E3214D6D CRC64;

Query Match
Best Local Similarity 3.9%; Score 22; DB 2; Length 538;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KVFVTGGRGGAGHAMVTVASDI 172
DB 152 KVFVTGGRGGAGHAMVTVASDI 173

RESULT 8
Q47014 PRELIMINARY; PRT; 538 AA.
AC Q47014;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Translocated intimin receptor tir (Translocated intimin co-receptor)
DE (EspE protein).
GN Name=tir; Synonym=espE;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC 84/110/1, and B65/56;
RA Krejany E.O.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=95ZG1;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (tir) of Shiga-toxicogenic Escherichia coli isolates belonging to serogroups O26, O111, and O157 react with sera from patients with hemolytic-uremic syndrome and exhibit marked sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=B10;
RX MEDLINE=20187493; PubMed=10722617;
RA DOI=10.1128/IAI.68.4.2171-2182.2000;
RA Marches O., Nougayrede J.P., Boullier S., Mainil J., Charlier G., Raymond I., Pohl P., Boury M., De Rycke J., Milon A., Oswald E.;
RT "Role of tir and intimin in the virulence of rabbit enteropathogenic Escherichia coli serotype O103:H2.";
RL Infect. Immun. 68:2171-2182 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=413/89-1;
RX MEDLINE=98294040; PubMed=9632251;
RA Deibel C., Kraemer S., Chakraborty T., Ebel P.;
RT "EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells where it appears as a tyrosine-phosphorylated 90 kDa protein.";
RL Mol. Microbiol. 28:463-474 (1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=413/89-1;
RX MEDLINE=98294040; PubMed=9632251;
RA Deibel C., Kraemer S., Chakraborty T., Ebel P.;
RT "EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells where it appears as a tyrosine-phosphorylated 90 kDa protein.";
RL Mol. Microbiol. 28:463-474 (1998).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=RDEC-1;
RX MEDLINE=97055784; PubMed=8900070; DOI=10.1016/0378-1097(96)00371-0;
RA Agin T.S., Canteley J.R., Boedeker E.C., Wolf M.K.;
RT "Characterization of the eaeA gene from rabbit enteropathogenic Escherichia coli strain RDEC-1 and comparison to other eaeA genes from bacteria that cause attaching-effacing lesions.";
RL FEMS Microbiol. Lett. 144:249-258 (1996).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=RDEC-1;
RX MEDLINE=21153569; PubMed=11254564;
RX DOI=10.1128/IAI.69.4.2107-2115.2001;
RA Zhu C., Agin T.S., Elliott S.J., Johnson L.A., Thate T.E., Kaper J.B., Boedeker E.C.;
RT "Complete nucleotide sequence and analysis of the locus of enterocyte
```

```

RT Effacement from rabbit diarrheagenic Escherichia coli RDEC-1."
RL Infect. Immun. 69:2107-2115(2001).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=83/39;
RX MEDLINE=22063667; PubMed=12067342;
RA Tauschek M., Strugnell R.A., Robins-Browne R.M.;
RT "Characterization and evidence of mobilization of the LEE
RT pathogenicity island of rabbit-specific strains of enteropathogenic
RT Escherichia coli.";
RL Mol. Microbiol. 44:1533-1550(2002).
DR EMBL; U59504; AAD19750.1; -.
DR EMBL; AF045568; AAC15683.1; -.
DR EMBL; AF200363; AAK26722.1; -.
DR EMBL; AF453441; AAL57549.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_receptor.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor.
SQ SEQUENCE 538 AA; 55410 MW; 3132A969B7B3D06C CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KVFVTGGRGGAGHAWTVASDI 172
DB 152 KVFVTGGRGGAGHAWTVASDI 173
|||||
|||||

RESULT 10
Q79C12 PRELIMINARY; PRT; 367 AA.
AC Q79C12;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Similar to the hypothetical 38.4kDa protein of REPEC 84/110/1; unknown
DE function (fragment).
OS Escherichia coli
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC RDEC-1;
RA Krejany E.O.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U59503; AAB02941.1; -.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_receptor.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 367 AA; 37797 MW; 00F84FCAC8FB5718 CRC64;

Query Match 3.6%; Score 20; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DLTKEAFQNPQKNVIDE 300
DB 102 DLTKEAFQNPQKNVIDE 121
|||||
|||||

```

```

RESULT 11
O50190 PRELIMINARY; PRT; 549 AA.
AC O50190;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RX MEDLINE=98050926; PubMed=9390560; DOI=10.1016/S0092-8674(00)80437-7;
RA Kenny B., Devinney R., Stein M., Reinscheid D.J., Frey E.A.;
RA Finlay B.B.;
RT "Enteropathogenic E. coli (EPEC) transfers its receptor for intimate
RT adherence into mammalian cells.";
RL Cell 91:511-520(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RA Stein M.S., Kenny B., Finlay B.B.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF013122; AAB88410.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_receptor.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor.
SQ SEQUENCE 549 AA; 56843 MW; 40C8B8B234409A08 CRC64;

Query Match 2.5%; Score 14; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IPPAPPLPSQTDGA 29
DB 16 IPPAPPLPSQTDGA 29
|||||
|||||

RESULT 12
O52147 PRELIMINARY; PRT; 550 AA.
AC O52147;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RX MEDLINE=98254123; PubMed=9593291;
RA Elliott S.J., Wainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
RA Lai L.C., Mchamara B.P., Donnenberg M.S., Kaper J.B.;
RT "The complete sequence of the locus of enterocyte effacement (LEE)
RT from enteropathogenic Escherichia coli E2348/69.";
RL Mol. Microbiol. 28:1-4(1998).
DR EMBL; AF022236; AAC38390.1; -.
DR HSSP; Q9KWH9; 1F02.

```

```

DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
SQ SEQUENCE 550 AA; 56509 MW; 19DD08A9BE9251CB CRC64;

Query Match          2.5%; Score 14; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IPPAPPLPSQTDGA 29
Db 16 IPPAPPLPSQTDGA 29

RESULT 13
ID O68258 PRELIMINARY; PRT; 551 AA.
AC O68258;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Translocated intimin receptor.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98187918; PubMed=9529069;
RA Voss E., Paton A.W., Manning P.A., Paton J.C.;
RT "Molecular analysis of Shiga toxinigenic Escherichia coli O111:H-
RT proteins which react with sera from patients with hemolytic-uremic
RT syndrome.";
RL Infect. Immun. 66:1467-1472(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (Tir) of Shiga-toxinigenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
DR EMBL; AF025311; AAC69249.1; -.
DR HSP; O9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
SQ SEQUENCE 551 AA; 56975 MW; 6EC95F76EF0F44CC CRC64;

Query Match          2.5%; Score 14; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IPPAPPLPSQTDGA 29
Db 16 IPPAPPLPSQTDGA 29

RESULT 14
ID O9KWH9 PRELIMINARY; PRT; 552 AA.
AC O9KWH9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
SQ SEQUENCE 552 AA; 57005 MW; A8D79EE22EE50A4B CRC64;

Query Match          2.5%; Score 14; DB 2; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IPPAPPLPSQTDGA 29
Db 16 IPPAPPLPSQTDGA 29

RESULT 15
ID O6GED0 PRELIMINARY; PRT; 163 AA.
AC O6GED0;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative exported protein.
GN OrderedLocusNames=SA2388;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282456;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Feil E.J., Lindey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; EX571856; CAG41369.1; -.
DR InterPro; IPR007921; CHAP.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 163 AA; 17007 MW; 202B20997858FCC1 CRC64;

Query Match          2.3%; Score 13; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 QTTTTTTTTTTTS 404
Db 40 QTTTTTTTTTTTS 52

```

Search completed: May 13, 2005, 12:06:49  
Job time : 203 secs

---



```

RESULT 1
US-09-946-374-243
/ Sequence 243, Application US/09946374
/ Publication No. US20030073129A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Hillan, Kenneth J.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: Secreted and Transferred
/ FILE REFERENCE: P2830P1C1
/ CURRENT APPLICATION NUMBER: US/09/946374
/ PRIOR FILING DATE: 2001-09-04
/ PRIOR APPLICATION NUMBER: 60/098716
/ PRIOR FILING DATE: 1998-09-01
/ PRIOR APPLICATION NUMBER: 60/098723
/ PRIOR FILING DATE: 1998-09-01
/ PRIOR APPLICATION NUMBER: 60/098749

```

[illegible]

1	APPLICANT:	Gerritsen, Mary E.
2	APPLICANT:	Godowski, Audrey
3	APPLICANT:	Godowski, Paul J.
4	APPLICANT:	Grimaldi, Christopher J.
5	APPLICANT:	Gurney, Austin L.
6	APPLICANT:	Watanabe, Colin K.
7	APPLICANT:	Wood, William I.
8	TITLE OF INVENTION:	SECRETED AND TRIM
9	TITLE OF INVENTION:	ACIDS ENCODING
10	FILE REFERENCE:	P3230R1C1
11	CURRENT APPLICATION NUMBER:	US/10/0
12	CURRENT FILING DATE:	2001-12-06
13	PRIOR APPLICATION NUMBER:	60/063435
14	PRIOR FILING DATE:	1997-10-29
15	PRIOR APPLICATION NUMBER:	60/064215
16	PRIOR FILING DATE:	1997-10-29
17	PRIOR APPLICATION NUMBER:	60/082797
18	PRIOR FILING DATE:	1998-04-22
19	PRIOR APPLICATION NUMBER:	60/083495
20	PRIOR FILING DATE:	1998-04-29
21	PRIOR APPLICATION NUMBER:	60/085579
22	PRIOR FILING DATE:	1998-05-15
23	PRIOR APPLICATION NUMBER:	60/087759
24	PRIOR FILING DATE:	1998-06-02
25	PRIOR APPLICATION NUMBER:	60/088021
26	PRIOR FILING DATE:	1998-06-04
27	PRIOR APPLICATION NUMBER:	60/088029
28	PRIOR FILING DATE:	1998-06-04
29	PRIOR APPLICATION NUMBER:	60/088030
30	PRIOR FILING DATE:	1998-06-04
31	PRIOR APPLICATION NUMBER:	60/088734
32	PRIOR FILING DATE:	1998-06-10
33	PRIOR APPLICATION NUMBER:	60/088740
34	PRIOR FILING DATE:	1998-06-10
35	PRIOR APPLICATION NUMBER:	60/088811
36	PRIOR FILING DATE:	1998-06-10
37	PRIOR APPLICATION NUMBER:	60/088824
38	PRIOR FILING DATE:	1998-06-10
39	PRIOR APPLICATION NUMBER:	60/088825
40	PRIOR FILING DATE:	1998-06-10
41	PRIOR APPLICATION NUMBER:	60/088863
42	PRIOR FILING DATE:	1998-06-11
43	PRIOR APPLICATION NUMBER:	60/089105
44	PRIOR FILING DATE:	1998-06-12
45	PRIOR APPLICATION NUMBER:	60/089514
46	PRIOR FILING DATE:	1998-06-16
47	PRIOR APPLICATION NUMBER:	60/089653
48	PRIOR FILING DATE:	1998-06-17
49	PRIOR APPLICATION NUMBER:	60/089952
50	PRIOR FILING DATE:	1998-06-19
51	PRIOR APPLICATION NUMBER:	60/090246
52	PRIOR FILING DATE:	1998-06-22
53	PRIOR APPLICATION NUMBER:	60/090444
54	PRIOR FILING DATE:	1998-06-24
55	PRIOR APPLICATION NUMBER:	60/090688
56	PRIOR FILING DATE:	1998-06-25
57	PRIOR APPLICATION NUMBER:	60/090696
58	PRIOR FILING DATE:	1998-06-25
59	PRIOR APPLICATION NUMBER:	60/090862
60	PRIOR FILING DATE:	1998-06-26
61	PRIOR APPLICATION NUMBER:	60/091628
62	PRIOR FILING DATE:	1998-07-02
63	PRIOR APPLICATION NUMBER:	60/096012
64	PRIOR FILING DATE:	1998-08-10
65	PRIOR APPLICATION NUMBER:	60/096757
66	PRIOR FILING DATE:	1998-08-17
67	PRIOR APPLICATION NUMBER:	60/096949
68	PRIOR FILING DATE:	1998-08-18
69	PRIOR APPLICATION NUMBER:	60/096959
70	PRIOR FILING DATE:	1998-08-18
71	PRIOR APPLICATION NUMBER:	60/097954
72	PRIOR FILING DATE:	1998-08-26
73	PRIOR APPLICATION NUMBER:	60/097971

Best Local Similarity 19.2%; Pred. No. 0.0021;  
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

RESULT 2  
US-10-006-867-100  
; Sequence 100, Application US/10006867  
; Publication No. US20020119130A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton,Dan L.  
; APPLICANT: Filvaroff,Ellen

;  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/097979  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/098749  
; PRIOR FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: 60/099741  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099763  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099792  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099812  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099815  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100627  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100662  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100683  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/100684  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/100930  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/101279  
; PRIOR FILING DATE: 1998-09-22  
; PRIOR APPLICATION NUMBER: 60/101475  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101738  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101743  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101916  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/102570  
; PRIOR FILING DATE: 1998-09-30  
; PRIOR APPLICATION NUMBER: 60/103449  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: 60/103678  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 60/103679  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 60/103711  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 60/105000  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 60/105002  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 60/105881  
; PRIOR FILING DATE: 1998-10-27  
; PRIOR APPLICATION NUMBER: 60/106030  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: 60/106464  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 60/106856  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: 60/108807  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: 60/112419  
; PRIOR FILING DATE: 1998-12-15  
; PRIOR APPLICATION NUMBER: 60/112422  
; PRIOR FILING DATE: 1998-12-15  
; PRIOR APPLICATION NUMBER: 60/112853  
; PRIOR FILING DATE: 1998-12-16  
; PRIOR APPLICATION NUMBER: 60/113011  
; PRIOR FILING DATE: 1998-12-16  
; PRIOR APPLICATION NUMBER: 60/112854  
; PRIOR FILING DATE: 1998-12-16  
; PRIOR APPLICATION NUMBER: 60/113300  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/113408  
; PRIOR FILING DATE: 1998-12-22

;  
; PRIOR APPLICATION NUMBER: 60/113430  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113621  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/114223  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 60/115614  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/116527  
; PRIOR FILING DATE: 1999-01-20  
; PRIOR APPLICATION NUMBER: 60/116843  
; PRIOR FILING DATE: 1999-01-22  
; PRIOR APPLICATION NUMBER: 60/119285  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: 60/119287  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: 60/119525  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/119549  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/120014  
; PRIOR FILING DATE: 1999-02-11  
; PRIOR APPLICATION NUMBER: 60/129122  
; PRIOR FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: 60/129674  
; PRIOR FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/131291  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/138387  
; PRIOR FILING DATE: 1999-06-09  
; PRIOR APPLICATION NUMBER: 60/144791  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/169495  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/175481  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 60/191007  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/199397  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/380139  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 09/311832  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380138  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380142

Query Match 6.3%; Score 180; DB 13; Length 596;

Best Local Similarity 19.2%; Pred. No. 0.0021;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

Qy 4 GNLGHNPNVNSIPPAPPLPSQTDGAGRGQLINSTGLSGRALPTVRNSWADSGDNRA 63

Db 13 GLLHLLEAATNS-----NETSTANTSGSVISSG-----ASTATNSG 49

Qy 64 SDVPGLPVNPNRLAASEITLNDGFEVLHDGFLDTLNRQIGSSV-FRVETQEDGKHIAGV 122

Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSESTASSGSIATN 107

Qy 123 QRNGVETS--VVLSDQEVARLQSIDPECKDFVFTGRRGGAGHAMVTVASDITEARQIRL 180

Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVTHSGSVTSSGASTATNSESTVSSRAS 163

Qy 181 ELLEPKGTGSGKAGESKVGELRESNSGAENTTETQTSTSSLRSDPKLWALGTVAT 240

Db 164 TATNSESSLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211

Qy 241 GLIGLAATGIQVALTPEPDSPTTDDAASATEATATRDQLTKEAFQNPQNKVNIDE 300

Db 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264

QY 301 LG-----NAIPSGVLKDDVVANIEQAKAAGEAKQQAIIENNAQAKKYDQQAQR 351  
DB 265 SGASTATNSDSTVSSGA---STATNSESTTSSGAST-----ATN 302  
QY 352 QBELKVSSGAGYGL-SGALILGGIGGVAVTAALHRKNQPVQETTTTTTTTTTTSARTVEN 410  
DB 303 SSSSTSSGASTATNSDSTSSGAGTATNSESTVSSGISTVTNSESTPSSGANTATN 362  
QY 411 KPANNTPAGNVDTFGSEPTMESRRSSMASTSTFTDTSSTIGTVQNPYADVKT----- 464  
DB 363 SSSSTSSGANTAT-NSBSTVSSGASTATNSESTTSSGVTSTATNSESTTSSGASTAT 421  
QY 465 LHDSQVPTNSNTSVQNMGTDSVYVSTIQHPRTTDTNGALLGNPSAGIOSTVARIAL 524  
DB 422 NSDSTSSSTSEASTATNSESTVSSGISTVTNSESTTSSGANTATNSGSSVTS----- 474  
QY 525 SGGLRHDGMLTGCNGNSAVNTSNNPPAPG 553  
DB 475 AGSGTAALTGMTTSHSA-STAVSEAKPG 502

## RESULT 3

US-10-052-586-310  
; Sequence 310, Application US/10052586  
; Publication No. US20020127584A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C1  
; CURRENT APPLICATION NUMBER: US/10/052,586  
; CURRENT FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063564  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063734  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/063870  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066120  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/066466  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/066772  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/069335  
; PRIOR FILING DATE: 1997-12-11  
; PRIOR APPLICATION NUMBER: 60/069425  
; PRIOR FILING DATE: 1997-12-12  
; PRIOR APPLICATION NUMBER: 60/069870  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/068017  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/083495  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083496  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083499  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083559  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/084366  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579

;  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/086023  
; PRIOR FILING DATE: 1998-05-18  
; PRIOR APPLICATION NUMBER: 60/086392  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/086486  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087098  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087208  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088722  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088740  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088811  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088825  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088863  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089090  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17

;  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089908  
  
Query Match 6.3%; Score 180; DB 13; Length 596;  
Best Local Similarity 19.2%; Pred. No. 0.0021;  
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;  
  
QY 4 GNLGHNPNVNSIPPAPPLPSQTDGAGRGGLINSTGLGRALFTPVNSMADSGDNRA 63  
Db 13 GLLHLEAATNS-----NETSTANTSGSVISSG-----ASTATNSG 49  
QY 64 SDVPCLPUNPMRLAASEITLNDGPEVLHDHGLDPLTLNQIGSSV-PRVETQEDGKHIAVG 122  
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSESTASGSIATN 107  
QY 123 QRNGVETS--VVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRL 180  
Db 108 SESSTTSSGASTATNSE---SSTPSSGASTVTVNSGSSVTSSGASTATNSESTVSSRAS 163  
QY 181 ELLBPKGTGSGKAGESKVGELRESNGAENTTETOTSTSTSLRSDPKLWALGTVAT 240  
Db 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211  
QY 241 GLIGLAATGIYOALATPEPDSPTTDDPDAASATETATRDQLTKEAFQNPDPNQKVNI 300  
Db 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTRTNSGAGTATNSESTTS 264  
QY 301 LG-----NAIPSGVLKDDVVANIBEQAKAAGEAKQAIAENNAQAQKYDEQQAQR 351  
Db 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302  
QY 352 QEELKVSAGGYGL-SGALIILGGIGVAVTAALHRKNQPVQOTTTTTTTTTTSTARTVEN 410  
Db 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTTVSSGISTVTVNSESTTSSGASTATN 362  
QY 411 KPAANTPAQGNVDTPGSEDTMESRRSSMASTSSFFDTSSTGTVQNPVADKTS----- 464  
Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESTTSSGASTATNSESTTSSGASTAT 421  
QY 465 LHDSQVTSNSTSVQNMGNSTDVVYSTIQHPPRDITDNGARLLGNPAGIQSTYARLAL 524  
Db 422 NSDSSTTSSSEASTATNSESTTVSSGISTVTVNSESTTSSGASTATNSGSSVTS----- 474  
QY 525 SGGLRHDHMGGLTGGNSGAVNTSNPPAPG 553  
Db 475 AGSGTAALTGMHTTSHGA-STAVSEAKPG 502

RESULT 4  
US-10-063-547-100  
; Sequence 100, Application US/10063547  
; Publication No. US20020182638A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3230R1C1  
; CURRENT APPLICATION NUMBER: US/10/063,547  
; PRIOR FILING DATE: 2002-05-02  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 100

LENGTH: 596  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-063-551-100

Query Match 6.3%; Score 180; DB 13; Length 596;  
Best Local Similarity 19.2%; Pred. No. 0.0021;  
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGPGSRALFTPPVNSMADSGDNRA 63  
DB 13 GLLHLEAATNS-----NETSTANTGSSVISG-----ASTATNSG 49  
QY 64 SDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDLNROIGSSV-PRVETQEDGKHIAVG 122  
DB 50 SVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEPFTASSGISAIATN 107  
QY 123 QNNGVETS--VVLSDQYARLQSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQIRL 180  
DB 108 SESSITSSGASTATNSE-----SSTPSSGASTVTVNSGSSVTSSGASTATNSESTVSSRAS 163  
QY 181 ELLEPKGTGSGKAGESKVGELRESNGAENTTQTSTSTSSLRSDPKMLALGTVAT 240  
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211  
QY 241 GLIGLAATGIQALALTPEPDSPTTTPDPAASATETATRDQLTKEAFQNPQKNVIDE 300  
DB 212 N-----SESTVSSRATATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264  
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQQAIIENNAQAQKDYDEQAKR 351  
DB 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302  
QY 352 QBELKVSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTSTARTVEN 410  
DB 303 SESSITSSGASTATNSDSSTTSSGASTATNSESTVSSGISTVTVNSSESTPSSGANTATN 362  
QY 411 KPAANTPAQGNVDTPEGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADVITS-----464  
DB 363 SESSITSSGANTAT-NSESTVSSGASTATNSESTTSSGVTATNSESTTSSGASTAT 421  
QY 465 LHDSQVPTNSNTSVQNMGNNTDSVVYSTIQHPRTDTTNGARLLGNPSAGIOSTVARLAL 524  
DB 422 NDSSTTSSGASTATNSESTVSSGISTVTVNSSESTTSSGANTATNSGSSVTS-----474  
QY 525 SGGLRDMGGLTGGNSAVNTSNPPAPG 553  
DB 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 5  
US-10-063-551-100  
; Sequence 100, Application US/10063551  
; Publication No. US20020183494A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3230R1C1  
; CURRENT APPLICATION NUMBER: US/10/063,551  
; PRIOR FILING DATE: 2002-05-02  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 100  
; LENGTH: 596

TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-063-551-100

Query Match 6.3%; Score 180; DB 13; Length 596;  
Best Local Similarity 19.2%; Pred. No. 0.0021;  
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGPGSRALFTPPVNSMADSGDNRA 63  
DB 13 GLLHLEAATNS-----NETSTANTGSSVISG-----ASTATNSG 49  
QY 64 SDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDLNROIGSSV-PRVETQEDGKHIAVG 122  
DB 50 SVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEPFTASSGISAIATN 107  
QY 123 QNNGVETS--VVLSDQYARLQSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQIRL 180  
DB 108 SESSITSSGASTATNSE-----SSTPSSGASTVTVNSGSSVTSSGASTATNSESTVSSRAS 163  
QY 181 ELLEPKGTGSGKAGESKVGELRESNGAENTTQTSTSTSSLRSDPKMLALGTVAT 240  
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211  
QY 241 GLIGLAATGIQALALTPEPDSPTTTPDPAASATETATRDQLTKEAFQNPQKNVIDE 300  
DB 212 N-----SESTVSSRATATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264  
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQQAIIENNAQAQKDYDEQAKR 351  
DB 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302  
QY 352 QBELKVSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTSTARTVEN 410  
DB 303 SESSITSSGASTATNSDSSTTSSGASTATNSESTVSSGISTVTVNSSESTPSSGANTATN 362  
QY 411 KPAANTPAQGNVDTPEGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADVITS-----464  
DB 363 SESSITSSGANTAT-NSESTVSSGASTATNSESTTSSGVTATNSESTTSSGASTAT 421  
QY 465 LHDSQVPTNSNTSVQNMGNNTDSVVYSTIQHPRTDTTNGARLLGNPSAGIOSTVARLAL 524  
DB 422 NDSSTTSSGASTATNSESTVSSGISTVTVNSSESTTSSGANTATNSGSSVTS-----474  
QY 525 SGGLRDMGGLTGGNSAVNTSNPPAPG 553  
DB 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 6  
US-10-174-590-310  
; Sequence 310, Application US/10174590  
; Publication No. US20030008352A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C42  
; CURRENT APPLICATION NUMBER: US/10/174,590  
; CURRENT FILING DATE: 2002-06-18  
; Prior application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 310



```
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-310

Query Match      6.3%; Score 180; DB 14; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGPLGRALFTPVNRNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLRAASEITLNDGFVLDHGHPLDNLNRQIGSSV-FRVETQEDGKHIAVG 122
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSGISTATNSESTASSGISIATN 107
QY 123 QRNGVETS--VVLSDQEVYARLQSIDPEGKQKVFVTGGRGGAGHAMVTVASDITEARQIRL 180
Db 108 SESSTTSSGASTATNSE---SSTPSSGASTVTNNGSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGESKGAGESKGVGELRESNGAENTTETQTSTSSLRSDPKLWLALGTVAT 240
Db 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIQOALALTPEPDSPTTDPDAAASATETATRDQLTKEAFQNPQNKVNIDE 300
Db 212 N-----SESVTSRSTATNSESTT---SSGASTATNSESTTSSGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIIENNAQAQKYDQQAQR 351
Db 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QBELKVSAGYGL-SGALLGGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVEN 410
Db 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTVSSGISTVTNSESTPSSGANTATN 362
QY 411 KPANTPAQGNVDTPGSEDPTMESRRSSMASTSTFTDTSIGTVQNPVADVKT-----464
Db 363 SESSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGVTATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNNTDSVVYSTIQHPRTDNGARLLGNPSAGIQSTYARLAL 524
Db 422 NSDSTTSSSEASTATNSESTVSSGISTVTNSESTTSSGANTATNSESTTSSGASTATN 474
QY 525 SGGLRHDHMGGLTGGNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 7
US-10-176-758-310
; Sequence 310, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-310

Query Match      6.3%; Score 180; DB 14; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGPLGRALFTPVNRNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLRAASEITLNDGFVLDHGHPLDNLNRQIGSSV-FRVETQEDGKHIAVG 122
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSGISTATNSESTASSGISIATN 107
QY 123 QRNGVETS--VVLSDQEVYARLQSIDPEGKQKVFVTGGRGGAGHAMVTVASDITEARQIRL 180
Db 108 SESSTTSSGASTATNSE---SSTPSSGASTVTNNGSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGESKGAGESKGVGELRESNGAENTTETQTSTSSLRSDPKLWLALGTVAT 240
Db 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIQOALALTPEPDSPTTDPDAAASATETATRDQLTKEAFQNPQNKVNIDE 300
Db 212 N-----SESVTSRSTATNSESTT---SSGASTATNSESTTSSGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIIENNAQAQKYDQQAQR 351
Db 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QBELKVSAGYGL-SGALLGGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVEN 410
Db 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTVSSGISTVTNSESTPSSGANTATN 362
QY 411 KPANTPAQGNVDTPGSEDPTMESRRSSMASTSTFTDTSIGTVQNPVADVKT-----464
Db 363 SESSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGVTATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNNTDSVVYSTIQHPRTDNGARLLGNPSAGIQSTYARLAL 524
Db 422 NSDSTTSSSEASTATNSESTVSSGISTVTNSESTTSSGANTATNSESTTSSGASTATN 474
QY 525 SGGLRHDHMGGLTGGNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 8
US-10-175-737-310
; Sequence 310, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
```

```
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-310

Query Match
  6.3%; Score 180; DB 14; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGLGRALFTPVNRNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLAASEITLNDGFEVLHDHGDPLDNLNQIGSSV-FRVETQEDGKHIAVG 122
Db 50 SSVTSSGVSATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEFSTASSGSIATN 107
QY 123 QRNQVETS--VVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRL 180
Db 108 SESSTSSGASTATNSE-----SSTPSSGASTVTNNGSSVTSNGASTATNSESTVSSRAS 163
QY 181 ELLPEKGTGESKGAGESKGVGELRESNGAENTTETQTSTSTSLRSDPKLWLALGTATVAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEPDSPTTDPDAASATETATRDQLTKEAFQNPQNKWNIDE 300
Db 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQAIENNAQAQKYDEQQAQR 351
Db 265 SGASTATNDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QBELKVVSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTISARTVEN 410
Db 303 SESSTTSSGASTATNDSSTTSSGAGTATNSESTVSSGISTVTNSESTPSSGANTATN 362
QY 411 KPANNTPAQGVNDTPGESDTHMESRRSSMASTSTFFDTSSIGTVQNPYADVKT-----464
Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESTTSSGVSATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGTDSVVYSTIOHPPTDNGARLLGNPAGIOSTVARIAL 524
Db 422 NDSSTTSSSEASTATNSESTVSSGISTVTNSESTTSSGANTATNSESTVSS-----474
QY 525 SGGLRHDMDGLTGGSNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 9
US-10-063-616-100
; Sequence 100, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,616
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170

; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-616-100

Query Match
  6.3%; Score 180; DB 14; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGLGRALFTPVNRNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLAASEITLNDGFEVLHDHGDPLDNLNQIGSSV-FRVETQEDGKHIAVG 122
Db 50 SSVTSSGVSATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEFSTASSGSIATN 107
QY 123 QRNQVETS--VVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRL 180
Db 108 SESSTSSGASTATNSE-----SSTPSSGASTVTNNGSSVTSNGASTATNSESTVSSRAS 163
QY 181 ELLPEKGTGESKGAGESKGVGELRESNGAENTTETQTSTSTSLRSDPKLWLALGTATVAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEPDSPTTDPDAASATETATRDQLTKEAFQNPQNKWNIDE 300
Db 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQAIENNAQAQKYDEQQAQR 351
Db 265 SGASTATNDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QBELKVVSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTISARTVEN 410
Db 303 SESSTTSSGASTATNDSSTTSSGAGTATNSESTVSSGISTVTNSESTPSSGANTATN 362
QY 411 KPANNTPAQGVNDTPGESDTHMESRRSSMASTSTFFDTSSIGTVQNPYADVKT-----464
Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESTTSSGVSATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGTDSVVYSTIOHPPTDNGARLLGNPAGIOSTVARIAL 524
Db 422 NDSSTTSSSEASTATNSESTVSSGISTVTNSESTTSSGANTATNSESTVSS-----474
QY 525 SGGLRHDMDGLTGGSNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 10
US-10-174-581-310
; Sequence 310, Application US/10174581
; Publication No. US20030017540A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deanovers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C41
; CURRENT APPLICATION NUMBER: US/10/174,581
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 10/052586
```



Db	422	NSDSSTTSEASTATNSESSTVSSGISTVTNSESSTSSGANTATNSGSSVTS-----474
QY	525	SGGLRDMGGLTGGNSAVNTSNPPAPG 553
Db	475	AGSGTAALTGMHTTSHSA-STAVSEAKPG 502
RESULT 11		
US-10-176-483-310		
; Sequence 310, Application US/10176483		
; Publication No. US20030017541A1		
; GENERAL INFORMATION:		
; APPLICANT: Baker, Kevin P.		
; APPLICANT: Chen, Jian		
; APPLICANT: Desnoyers, Luc		
; APPLICANT: Goddard, Audrey		
; APPLICANT: Godowski, Paul J.		
; APPLICANT: Gurney, Austin L.		
; APPLICANT: Pan, James		
; APPLICANT: Smith, Victoria		
; APPLICANT: Watanabe, Colin K.		
; APPLICANT: Wood, William I.		
; APPLICANT: Zhang, Zemin		
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC		
; FILE REFERENCE: P3430RIC68		
; CURRENT APPLICATION NUMBER: US/10/176,483		
; CURRENT FILING DATE: 2002-06-20		
; Prior application removed - See File Wrapper or Palm		
; NUMBER OF SEQ ID NOS: 612		
; SEQ ID NO 310		
; LENGTH: 596		
; TYPE: PRT		
; ORGANISM: Homo Sapien		
US-10-176-483-310		
Query Match		
Best Local Similarity 6.3%; Score 180; DB 14; Length 596;		
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;		
QY	4	GNLGNPNVNNISIPAPPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVNNSMADSGDNRA 63
Db	13	GLLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY	64	SDVPLGNPNMRLAASEITLNDGFVLHDHGLDPLTNLRQIGSSV-FRVETQEDCKHIAVG 122
Db	50	SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEFSTASSGSIATN 107
QY	123	QRNGVETS--VVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQIL 180
Db	108	SESSTSSGASTATNSE-----SSTPSSGASTVTNNGSSSVTSSGASTATNSESSTVSSRAS 163
QY	181	ELLEPKGTGSGKAGESKGVGELRESNGAENTTTQTSTSLRSDPKLWALGTAT 240
Db	164	TATNSESSTSSGASTATN-SDSSTSSGASTATNSESSTSSG-----ASTAT 211
QY	241	GLIGLAATGIQVALATPEPDSPTTDPAAASATETATRDQLTKEAFQNPQKNVIDE 300
Db	212	N-----SESSTVSSRATATNSESSTT---SSGASTATNSESRTTNGAGTATNSESSTTS 264
QY	301	LG-----NAIPSGVLKDDVWANIIEQAKAAGEAKQQAIIENNAQAQKYDEQQAQR 351
Db	265	SGASTATNDSSTVSSGA---STATNSESSTSSGAST-----ATN 302
QY	352	QBELKVSGAGYGL-SGALLGGGIGVAVTAALHRKNQPVETTTTTTTTTTARTVEN 410
Db	303	SESSTSSGASTATNDSSTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATN 362
QY	411	KPANNTPAQNVDTTCGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADVKT-----464
Db	363	SESSTSSGANTAT-NSSESTVSSGASTATNSESSTSSGISTATNSESSTSSGASTAT 421
QY	465	LHDSQVPTNSNTSVQNMGNDSVVYSTIQHPPRDTTNGARLLGNPSAGIOSTVARLAL 524

Db 422 NSDSSTTSSEASTATNSSESTVSSGISTVTNSSESTTSSGANTATNSGSSVTS----- 474  
QY 525 SGGLRHDHMGGLTGGNSNAVNTSNPPAPG 553  
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

## RESULT 12

US-10-176-749-310  
; Sequence 310, Application US/10176749  
; Publication No. US20030017542A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C76  
; CURRENT APPLICATION NUMBER: US/10/176,749  
; CURRENT FILING DATE: 2002-06-20  
; Prior application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 310  
; LENGTH: 596  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-176-749-310

Query Match 6.3%; Score 180; DB 14; Length 596;  
Best Local Similarity 19.2%; Pred. No. 0.0021;  
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;  
QY 4 GNLGHNPNVNSIPPAPLPSTQDAGGRGQLINSTGPLGRALFTPVNRNSMADSGDNRA 63  
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49  
QY 64 SDVPGLPVNPMLAASEITLNDGFVLDHGHPLDTLNRQIGSSV-FRVETQEDGKHIAVG 122  
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSSEFSTASSGISIATN 107  
QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRI 180  
Db 108 SESSTSSGASTATNSE-----SSTPSSGASTVTNSGSSVTSSGASTATNSSESTVSSRAS 163  
QY 181 ELLPFGTGESKAGESKGVGELRESNGAENTTQTSTSTSSLRSDPKLWLALGTVAT 240  
Db 164 TATNSSESTLSSGASTATN-SDSSTSSGASTATNSSESTSSG-----ASTAT 211  
QY 241 GLIGLAATGIQVLAALITPEPDSPTTDPDAASATETATRDOLTKEAFQNPQNKWNIDE 300  
Db 212 N-----SESSTVSSRASTATNSSESTT-----SSGASTATNSSESTTNGAGTATNSSESTTS 264  
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIIENNAQAOKKYDQQAQR 351  
Db 265 SGASTATNSDSSTVSSGA-----STATNSSESTSSGAST-----ATN 302  
QY 352 QEELKVSAGAGYGL-SGALILGGGIGVAVTAALHRKNQFVEQTTTTTTTTTTSARTVEN 410  
Db 303 SESSTSSGASTATNSDSSTSSGASTATNSSESTVSSGISTVTNSSESTPSSGANTATN 362  
QY 411 KPANNTPAQGVNDTPGSEDTMESRRSSMASSTSTFTDSSIGTGVQNPYADVKT----- 464  
Db 363 SESSTSSGANTAT-NSSESTVSSGASTATNSSESTTSSGVTSTATNSSESTTSSGASTAT 421

QY 465 LHDSQVPTNSNTSVQNMGNNTDSVVYVSTIQHPRPDTTNGARLLGNPSPAGIQSTVARLAL 524  
Db 422 NSDSSTTSSEASTATNSSESTVSSGISTVTNSSESTTSSGANTATNSGSSVTS----- 474  
QY 525 SGGLRHDHMGGLTGGNSNAVNTSNPPAPG 553  
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

## RESULT 13

US-10-176-914-310  
; Sequence 310, Application US/10176914  
; Publication No. US20030017543A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C83  
; CURRENT APPLICATION NUMBER: US/10/176,914  
; CURRENT FILING DATE: 2002-06-20  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 310  
; LENGTH: 596  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-176-914-310

Query Match 6.3%; Score 180; DB 14; Length 596;  
Best Local Similarity 19.2%; Pred. No. 0.0021;  
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;  
QY 4 GNLGHNPNVNSIPPAPLPSTQDAGGRGQLINSTGPLGRALFTPVNRNSMADSGDNRA 63  
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49  
QY 64 SDVPGLPVNPMLAASEITLNDGFVLDHGHPLDTLNRQIGSSV-FRVETQEDGKHIAVG 122  
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSSEFSTASSGISIATN 107  
QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRI 180  
Db 108 SESSTSSGASTATNSE-----SSTPSSGASTVTNSGSSVTSSGASTATNSSESTVSSRAS 163  
QY 181 ELLPFGTGESKAGESKGVGELRESNGAENTTQTSTSTSSLRSDPKLWLALGTVAT 240  
Db 164 TATNSSESTLSSGASTATN-SDSSTSSGASTATNSSESTSSG-----ASTAT 211  
QY 241 GLIGLAATGIQVLAALITPEPDSPTTDPDAASATETATRDOLTKEAFQNPQNKWNIDE 300  
Db 212 N-----SESSTVSSRASTATNSSESTT-----SSGASTATNSSESTTNGAGTATNSSESTTS 264  
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIIENNAQAOKKYDQQAQR 351  
Db 265 SGASTATNSDSSTVSSGA-----STATNSSESTSSGAST-----ATN 302  
QY 352 QEELKVSAGAGYGL-SGALILGGGIGVAVTAALHRKNQFVEQTTTTTTTTTTSARTVEN 410  
Db 303 SESSTSSGASTATNSDSSTSSGASTATNSSESTVSSGISTVTNSSESTPSSGANTATN 362  
QY 411 KPANNTPAQGVNDTPGSEDTMESRRSSMASSTSTFTDSSIGTGVQNPYADVKT----- 464  
Db 363 SESSTSSGANTAT-NSSESTVSSGASTATNSSESTTSSGVTSTATNSSESTTSSGASTAT 421

Query Match	6.3%;	Score 180;	DB 14;	Length 596;
Best Local Similarity	19.2%;	Pred. No. 0.0021;		
Matches	109;	Conservative 90;	Mismatches 272;	Indels 98; Gaps 19;
QY	4	GNLGHNEVNVNSIPPAPPLPSQTDGAGRGQLINSTGLGSRALFTVPRNSMADSGDRA	63	
DB	13	GLLHLLEAATNS-----NETSTANTGSSVISSG-----ASTATNSG	49	
QY	64	SDVPGLEVPNPMRLAASEITLNDGFEVLHDHGPDLTLNRQIGSSV-FRVETQEDCKHIAVG	128	
DB	50	SSVTSSGCVSTATISGSSVTSN-GVSIIV-TNSEFHTTSSGISTATNSSEFSTASSGISATIN	107	
QY	123	QRNGVETS--VVLSDQBYARLQSIDPGKOKFVFTGGRGGAGHAMVTVASDITEARQIL	180	
DB	108	SESSTTSSGASTATNSE----SSTPSSGASVTNNGSVTSSTSGASTATNSSESSTVSSRAS	163	
QY	181	ELLEPKGTGSGKAGESKGVCGLRESNGSAGENTTETOTSTSTSLRSDPKLWLALGTVAT	240	
DB	164	TATNSESTTSSGASTATN-SDSSTTSSGASTATNSSESTTSSG-----ASTAT	211	
QY	241	GLICLAATGIVQALATPEPDSPTTTDPDAAAATETATRDQLTKFAQNPNDKNKVIDE	300	
DB	212	N-----SESSTVSSRASATNSSESTT---SSGASTATNSSESTTNGAGTATNSSESTTS	264	
QY	301	LG-----NAIPSGVLKDDVVANIEBOAKAAGEAKQQAENNAQAOKKYDEQQAQR	351	
DB	265	SGASTATNSDSSTVSSGA---STATNSSESTTSSGAST-----ATN	302	
QY	352	QEBELKVSSSGAGYGL-SCALILGGIGIVAVTAALHRKNQPVQETTTTTTTTTTTSARTVEN	410	
DB	303	SESSTTSSGASTATNSDSSTTSSGAGTATNSSESTVSSGGISTVTNSESSTPSSGANTATN	362	
QY	411	KPANNTPAQGNVTPGSEDWTMESRRSSMASTSTSTFTTSSIGTVQNFPYADVKTST-----	464	

Query Match	6.3%	Score 180;	DB 14;	Length 596;
Best Local Similarity	19.2%	Pred. No. 0.0021;		
Matches 109;	Conservative 90;	Mismatches 272;	Indels 98;	Gaps 19;
QY	4	GNLGHNPVNNSIPPAPPLPSQTDGAGGRQLNSTCPLGSRALFTVPRNSMADSGDNRA	63	
DB	13	GLLLHLEAATNS-----NETSTANTSGSSVSSG-----ASTATNSG	49	
QY	64	SDVPGPLVPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSV-PRVETQEDCKHIAVG	122	
DB	50	SSVTSRGVSTATISGSSVTN-GVSIIV-TNSEFHTTSSGISTATNSSEFSTASSGISATIN	107	
QY	123	QRNGVETS--VILSDOEYARLQIDPEGKDKFVFTGRCGAGHAMVTVASDITEARQIL	180	
DB	108	SESSITSSGASTATNSE---SSTPSSGASTVTVNSGSSVTSAGASTATNSSESVTSGRAS	163	
QY	181	ELLEPKGTGSGKAGESKGVGELRESNGAENTTETQTSTSTSLRSDPKLWLALGTVAI	240	
DB	164	TATNSSESTLSSGASTATN-SDSSTTSSGASTATNSSESTTSSG-----ASTAT	211	
QY	241	GLIGLAATGIVQALALTPEPDSFTTTDDPAAASATATATRDQLTKEAFQPNPDKNQVIDE	300	
DB	212	N-----SESTVSSRASATATNSSESTT---SSGASTATNSSESTTSSGAGTATNSSESTTS	264	
QY	301	LG-----NAIPSGVLKDDVVAITEQAKAAGEAKQAIENNAQAOKKYDEQOAKR	351	
DB	265	SGASTATNSDSTSTVSSGA---STATNSSESTTSSGAST-----ATN	302	
QY	352	QBELKVSSGAGYGL-SCALILGGGIGGVAVTAALHRKNQPVQEQTITTTTTTTTSARTVEN	410	
DB	303	SESSITSSGASTATNSDSTTSSGAGTATNSSESTVSSGISTVTVNSSESTPSSGANTATIN	362	
QY	411	KPANINTPAQNVDTPGSEBDTMSERRRSMASTSTPFDTSSIGTVQNPIYADVKTSS	464	
DB	363	SESSITSSGANTAT-NSESTVSSGASTATNSSESTTSSGCVSTATNSSESTTSSGASTAT	421	





Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	172	6.1	2137	3	US-09-134-001C-4463	Sequence 4463, App
2	172	6.1	2736	4	US-09-252-991A-30227	Sequence 30227, A
3	170	6.0	529	4	US-09-248-796A-16703	Sequence 16703, A
4	166	5.8	2870	4	US-09-479-467A-15	Sequence 15, Appl
5	166	5.8	3178	4	US-09-479-467A-4	Sequence 4, Appl
6	163.5	5.8	821	4	US-09-556-877-195	Sequence 195, App
7	163.5	5.8	821	4	US-09-620-412C-195	Sequence 195, App
8	163.5	5.8	821	4	US-09-598-419-195	Sequence 195, App
9	163.5	5.8	1776	4	US-09-556-877-179	Sequence 179, App
10	163.5	5.8	1776	4	US-09-620-412C-179	Sequence 179, App
11	163.5	5.8	1776	4	US-09-598-419-179	Sequence 179, App
12	159.5	5.6	441	4	US-09-248-796A-15936	Sequence 15936, A
13	158	5.6	556	4	US-09-248-796A-22338	Sequence 22338, A
14	157.5	5.5	827	4	US-09-248-796A-17307	Sequence 17307, A
15	152.5	5.4	2045	4	US-09-949-016-10491	Sequence 10491, A
16	152	5.4	2314	3	US-09-268-347-49	Sequence 49, Appl
17	151	5.3	1306	4	US-09-538-092-330	Sequence 330, Appl
18	149.5	5.3	785	4	US-09-538-092-872	Sequence 872, App
19	148	5.2	2035	1	US-08-046-585-5	Sequence 5, Appl
20	148	5.2	2035	1	US-08-393-703-5	Sequence 5, Appl
21	148	5.2	2035	5	PCR-US93-11721-5	Sequence 5, Appl
22	148	5.2	2142	4	US-09-540-236-3459	Sequence 3459, Ap
23	147	5.2	525	4	US-09-107-532A-5095	Sequence 5095, Ap
24	147	5.2	985	5	PCR-US96-03916-6	Sequence 6, Appl
25	147	5.2	985	5	PCR-US96-03916-66	Sequence 66, Appl
26	146.5	5.2	1095	5	US-09-206-942-69	Sequence 69, Appl
27	146.5	5.2	1536	1	US-08-038-682-2	Sequence 2, Appl

QY 344 -----YDEQAKQKQELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQPVQETTT- 395  
Db 1125 ESLSTSVSDSTASSTSESEGNASASTSLGSLSTISDSSTSTSDSASTSTS 1184  
QY 396 -----TTTTTTTARTVENKANNTPAQGNVDTFGSEDTMESRSSMASTSTSTFFDTS 449  
Db 1185 ESESSTSLSESTSLSDSTSTSESASTSTSESASTSTSES-TSLSESTSTSVSDST 1243  
QY 450 SIGTVQNPYADVKTSLHDSQ-VPTGNS---NTSVQNMGT---DSVYSTTQHPPRDTTD 502  
Db 1244 SASTSDS---ASTSTSVSDSESASTSISELSSTSVSDSTSTSDSASTSTSE---SDSTS 1298  
QY 503 NGARLLGNPSAGI-----QSTYARLALSGGURHDMGGLTGGSNS-A 542  
Db 1299 ESTLSSESISTSVSDSTGASTSDSASTSTSESESDSASTSLSGSTSTSLSDSTSTSDS 1358  
QY 543 VNTSNN 548  
Db 1359 ASTSTS 1364  
  
RESULT 2  
US-09-252-991A-30227  
; Sequence 30227, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30227  
; LENGTH: 2736  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30227

Query Match 6.1%; Score 172; DB 4; Length 2736;  
Best Local Similarity 23.9%; Pred. No. 0.0045;  
Matches 148; Conservative 60; Mismatches 242; Indels 170; Gaps 33;  
  
QY 2 PIGNLGHNPVNNSTIPAPLPSQTDGAGRGQLNSTGPLGSRALFTPVNRNSMADSDN 61  
Db 723 PIGQVTADSGSNWSTPTTLPN-----GTVVNATATDAS-----GNTSAGSVT 767  
QY 62 RASDVPGLPV-NPMRLAASEITLNDGFVLHDHGFPLDLNRQIGSSVFRVETQEDGKHIA 120  
Db 768 VDSVAPATPVINP-----SNGTTLSTGTAEP-----GSSV-----TLTDGNGNP 805  
QY 121 VGORN-----GVETSVVLSDOEVARLQSIDPEGDKFVFTGGRGA---GHAMVTVAS 170  
Db 806 IQQVTADSGSNWSTPTPLADGTVVNATATDPAEN-----TSGGSGTVDGVAPTFTTV 860  
QY 171 DITEARQRIELLEPKGTGESKGA-----GESKGVGELRENSGAENTTETQTSTSSSL 225  
Db 861 NLSNGSSL-----SGTAEPGSTVILTNGNGNPIAEVTDGSG--NWYTT----- 902  
QY 226 RSDPKLWALGTAVTGLGLAATGIVQALALTPEPDSPT-----TTDPPAA 271  
Db 903 ---PSTPIANGTVVNVVQAADAGNSPGASVTVDSPAAPVVPVNSGTTLSGTAEPGAT 959  
QY 272 ASATETATRDQLTKEAFQNPQNKVNIDELGN--AIPSGVLKDDVVVNIIEQAKAAGEA 329  
Db 960 VSLT-----DNGNPTICQ-VTADSGSNWSTPTGTPLANGTVVNVATASDPTGNTSA 1008  
QY 330 -KQQAIIENNAQKQYDEQAKQKQELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQ 388

Db 1009 PASTTVDSVAPAAPVVPNSNG-----VVISGTAEPGATVTLTDGSG-----N 1050  
QY 389 PVEQTTT-----TTTTTTTARTVEN---KPAANTPAQGN--VD--TPGSEDTMESR 434  
Db 1051 PIGQVTADSGSNWSTPTPLADGTVVNATATDPAAGTGGGSGTTVDAIAPATPVNLSN 1110  
QY 435 RSSMASTS---STPFDTSSIGTVQNPYADVKT--SLHDSQVP-TSNNSNTSVQNMGTDSV 488  
Db 1111 GSSLSGTAEPGSTVILTNGNG---NPIAEVTDGSGNWTPTPPIANGTVVNVVVAEDAA 1167  
QY 489 VYSTIQHPPRDTT-----DNGARLLGNPSAGIQTSTYARLALSGGLRHDMDGL 535  
Db 1168 GNSS---PPATVTVDSAPPAPVNPNSGVVISGTAEGATVT---LTDAGG--NPICGV 1219  
QY 536 TG-GSNNAVNTSNPPAPGS 554  
Db 1220 TADSGSNWSTPTGTPPANGT 1239  
  
RESULT 3  
US-09-248-796A-16703  
; Sequence 16703, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 16703  
; LENGTH: 529  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-16703

Query Match 6.0%; Score 170; DB 4; Length 529;  
Best Local Similarity 23.0%; Pred. No. 5.7e-05;  
Matches 85; Conservative 47; Mismatches 150; Indels 88; Gaps 15;  
  
QY 205 ESNSGAENTTETQSTSTSSLSRSDPKLWALGTAVTGLGLAATGIVQALALTPEPDSPT 264  
Db 208 QSTSGKTTTTPVVTATSAITGNDN---TASTTDATG-----KTLTTVTSNDNTT 255  
QY 265 TTDPPAAASATETATRDQLTKEAFQNPQNKVNIDELGNALPISGVLKDDVVVNIIEQAK- 323  
Db 256 STGDDSTTASTGNDNTSTTTTATVKN-----NIDSTTNA-TDNTSKATTTGNDNTDSTK 309  
QY 324 -AAGEERAKQAIIENNAQ---AQKYDEQAKQKQELKVSSGAGYGLSGALILGGIGVAV 379  
Db 310 TASGTDSTTTTGDNNSESTTASTGNDN---NDSTTVTTGDDY-----TTV 351  
QY 380 TAALHRKNQVQETTTTTTTTTTISARTVENKPAANTPAQGNVD---TPGSEDTMESR 435  
Db 352 TT-----DNDNTASIVTTGNSNTTTLTDATGKOSTTTTNDGHDESTTVTTGGDDTATTKD 407  
QY 436 SSMASTSSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQNMGTDSVVYST--- 492  
Db 408 NNTASTTTKGDNDTATVNH-----SIESTVTTGDDNDTATGNDTASTTDTATGND 460  
QY 493 -----IQHPPRDTTDNGARLLGNPSAGIQTSTYARLALSGGLRHDMDGGLTGGSNSA 542  
Db 461 STTATTGNHSIESTPVTGDDNTATTGNDTAA--STVS-----TGNDNSA 503  
QY 543 VNTSNNPAP 552  
Db 504 SKTV--PPTP 511

**RESULT 4**

US-09-479-467A-15  
; Sequence 15, Application US/09479467A  
; Patent No. 6723557  
; GENERAL INFORMATION:  
; APPLICANT: Sternberg, Paul W.  
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE HOMOLOGS REQUIRED FOR MALE MATING BEHAVIOR IN NEMATODES AND ASSAYS BASED THEREON  
; FILE REFERENCE: 18021-2901B  
; CURRENT APPLICATION NUMBER: US/09/479,467A  
; CURRENT FILING DATE: 2000-01-06  
; PRIOR APPLICATION NUMBER: 60/115,127  
; PRIOR FILING DATE: 1999-01-06  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 2870  
; TYPE: PRT  
; ORGANISM: C. Elegans Lov-1 sy582 deletion protein  
**US-09-479-467A-15**

Query Match      5.8%; Score 166; DB 4; Length 2870;  
Best Local Similarity 20.3%; Pred. No. 0.0014;  
Matches 108; Conservative 89; Mismatches 223; Indels 112; Gaps 19;

Qy    68 GLPNPMLRLAASEITLND-----GFVLVHGHPLDLNLRIGSGSVFRVETQEDGKHI 119  
Db                 :|::||:||:||||:||||:||||:||||:||||:||||:||||:||||: 137 GLFLNSTWITLNEVNDDDEISIAVEAKYEVCYDDG----IDRCDSGLMW---LQVGGNEM 189  
Qy    120 A-VGORNGVTSLVDPEARLQSIDPECKDFVFTRGGRGGAHWTVVASDIPEARQR 178  
Db                 ||:||:||||:||||:||||:||||:||||:||||:||||:||||: 190 ALLGYREKCESGI--NEEYARMCRKPYSRK-----STAISDGQV 230  
Qy    179 ILELLPEPKGTGESKGAGESGVGEELRESNGAEENTTQTSTSSLSRSDPKLMALGTV 238  
Db                 |::|:||||:||||:||||:||||:||||:||||:||||:||||: 231 YYGQVILKVRAKFMSRTSGSPTLRMKRDAAGNTCDYTIESTSTSTTTPTTTTTTVTIV 290  
Qy    239 -ATGLLGLAATGIVALALTPEPDSPITT-----DPAAAASATETATRDQLTKFAEQ 289  
Db                 |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|: 291 TSTTVTPTSPTVTAMSTSTPSSTTIESTSTFTTASTSTGSTSTTQQOSSSTITS 350  
Qy    290 NPDNOKVINDELNAIPSGVKLD--DVVANIERQAAGA-EBAKKQAENNAQAOKKYDE 346  
Db                 ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|: 351 SPAS-----TTLSTSIPTTTTTPILTSSLSDPAICSYLDETTSFTTFMTLTSTTE 405  
Qy    347 QQAKROBELKVSQGAGYGLSCALILGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTT --S 404  
Db                 :|:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~ 406 EPSTSTTTTEVTS-----TSSPVTTTTEPTTLTSTASTSTEPS 445  
Qy    405 ARTVENKPA-----NNTPAQGVNDTCGSSEDTHESRRSSMASTSTSFDTSSIGTVQ 455  
Db                 ||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~ 446 TSITVVTSPTSPTVTSTVSSSSSTTTTPTSTESTSTSPSSSTTTTAPSTSTTGPPSS 505  
Qy    456 NPADVKTSLHDQVPSTNSGTSQNMGNTDSVVYSTIQHPPRDITDNGALLGNP---- 511  
Db                 :|:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~ 506 S-----SSTPSSASSVSSTASSQSSSTTQQSSSTTKSETTTSDGT----NPDPFYF 555  
Qy    512 -----SAGIOSTYARLASGLRHDMGGLTGCSSNAVNTSNPPPGSHRFV 558  
Db                 :|:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~ 556 VEKATTFFDYDSTSWNLINLSGL-----GIIGYQTSIBCTSI-----PTSSNVV 597

**RESULT 5**

US-09-479-467A-4  
; Sequence 4, Application US/09479467A  
; Patent No. 6723557  
; GENERAL INFORMATION:  
; APPLICANT: Sternberg, Paul W.  
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE HOMOLOGS REQUIRED FOR MALE MATING BEHAVIOR IN NEMATODES AND ASSAYS BASED THEREON

```

; FILE REFERENCE: 18021-2901B
; CURRENT APPLICATION NUMBER: US/09/479,467A
; CURRENT FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: 60/115,127
; PRIOR FILING DATE: 1999-01-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 3178
; TYPE: PRT
; ORGANISM: C. Elegans Lov-1 protein
US-09-479-467A-4

Query Match      5.8%; Score 166; DB 4; Length 3178;
Best local Similarity 20.3%; Pred. No. 0.0017;
Matches 108; Conservative 89; Mismatches 223; Indels 112; Gaps 19;

Qy 68 GLPNPRLAASITLND-----GFEVLHDHGLDPLTLNRQIGSSVFRVEQEDGKHI 119
Db 137 GLFLNTWITLNEVNDDEISIAVEAKYECYDDG-----IDRCGSLWW---LQVGNEM 189

Qy 120 A-VQORNGVTSVVLSDQEARLQSDPECKDFVFTGGRGGAGHAMVTVASDITEARQR 178
Db 190 ALIGYREKCESGRI--NEEYARMCKPYSEK-----STAISDSQGV 230

Qy 179 ILELLPEKGTGSKGAGESKVGGLRESNGAENTTQTSTSTSSLRSDPKLWALGTV 238
Db 231 YDGQVLKGVRAKQFSMRTSGSPTLRRMKRDAGDNTCDYTIESTSTSTTTPTTTTSTV 290

Qy 239 -ATGLIGLAATGIVQALALTPEDSPPTT-----DPDAASATATARDQLTKEAFQ 289
Db 291 TSTTTPSTSTVTTAMSTSTSTPTSTTSTTSTSTSTSTSTSTSTSTSTSTSTSTST 350

Qy 290 NPONQKNVIDELGNATPSGVKLD--DVVANIEQAKAAG--EAKQQAENNAQAQKKYDE 346
Db 351 SPSS-----TTLSTSPITTTTPTITSTLSSLPDNAICSYLDETTTSTTTMLTSTTE 405

Qy 347 QAKROBELKVSAGAGYGLSGLILGGGIGVAVTAALHRKNQPVQETTTTTTTTTT--S 404
Db 406 EPSTSTTTTTEVTS-----TSSTVTTTTEPTTLTSTASTSTSTSTSTSTSTSTST 445

Qy 405 ARTVENKPA-----NNTPAQGVNDTPGSEDTHESRRSSWASTSTSTFTDSSIGTVQ 455
Db 446 TSTVTTSPSTSPVTSTVTSSTSSSTTTTPTSTESTSTSPSTSTVTTSTTAPSTSTTGPSS 505

Qy 456 NPYADVKTSLHDQVPTNSNTSVQNMGNDSVVYSTIQHPPRDITDNGARLLGNP---- 511
Db 506 S-----SSTPSSVALSSSVSTASSQSSSTTQSSSTTKKSTTTSSDGT----NPDPYF 555

Qy 512 -----SAGIOSGYARLALSGLRHDMGGLTGGGNSAVNTSNPPAPGSHRFV 558
Db 556 VEXATTTFYDSTSNVLTNSGL-----GIIGYQTSICTS-----PTSSNVV 597

RESULT 6
US-09-556-877-195
; Sequence 195, Application US/095556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 195
; LENGTH: 821

```

```
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-195

Query Match
  5.8%; Score 163.5; DB 4; Length 821;
Best Local Similarity 21.9%; Pred. No. 0.00035;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;

Qy 79 SEITLNDGFEVLHDGPLDNLNRQIGSSVFRVETQED--GK----- 117
Db 276 TEQTKSNG-----NQDGSSETKDTQVSESPESPDDVLGKGGGIYTEKSLTITGITGTI 331
Qy 118 ----HIAVQQRNGVETSVVLSDQEVARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173
Db 332 DFVSNIAATDSGAGVFTKENLSCTNTNSLQFLKN-----SAGQHGGG-AVVTQMTSVT 382
Qy 174 EARQRIELLEPKGTGE-----SKGAGESKGVGELRESNGAENTTETOTSTSTSSLR 226
Db 383 NTTSE--SITTPPLVGEVIFSENTAKHGCGGICTNKLSLN-----LKTVTLTKNSAKE 434
Qy 227 SDPKLWLALGTVATGLIGLAATGIVQALALTPEDSPPTTDP----- 268
Db 435 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFPAS 480
Qy 269 ---DAAASATETATRDQL--TKEAFQNPQDNQKVNIDELGNAIPSGVLKDDVVANIEEQA 323
Db 481 TAEPAAAPSLTE-AESDQTDQDETSDTNSDID-VSIENILN-----VAINQNTSAK 528
Qy 324 AA-----GEEAKQQAENNAQAQKYDEQAQKQBELKVSSGAGYGLSGLALILG-----GG 374
Db 529 KGGAIYGGKAKLSRINN-----LELSGNSQDVGGGLCLITESVEFDA 570
Qy 375 IGVAVT-----AALHRKNQVPEQTTTTTTTTTTTTSARTVEN---KPANNTPAQG-- 420
Db 571 IGSLLSHVNSAAKEGGVHISKTVLNLKSTFTPADNTVKAIVESTPEAPEIIPVEGEE 630
Qy 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459
Db 631 STATENPNSNTEGSSANTNLEGSGQDTADTGTGVVNNESQDTSDTGNAESGEQLQDSTQS 690
Qy 460 DVKTSLHDSQVPTNSNTSVQNMGTDSVVTSTIQHPPRDTT---DNGARLLGNPSAGI 515
Db 691 NEENTLPNSSIDQSNENTDESSDSHTEITDESVSSSKSGSSTPDQDGAASSGAPS-GD 749
Qy 516 QSTYAR--LALSGGLRHDMGGLTGGSNSAVNTSNPPAPGS 554
Db 750 QSISANACLAKSAASTDSSPVSNSGSDVDTASSDNDPSSS 790

RESULT 7
US-09-620-412C-195
; Sequence 195, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 195
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-195

Query Match
  5.8%; Score 163.5; DB 4; Length 821;
Best Local Similarity 21.9%; Pred. No. 0.00035;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;

Qy 79 SEITLNDGFEVLHDGPLDNLNRQIGSSVFRVETQED--GK----- 117
```

```
Db 276 TEQTKSNG-----NQDGSSETKDTQVSESPESPDDVLGKGGGIYTEKSLTITGITGTI 331
Qy 118 ----HIAVQQRNGVETSVVLSDQEVARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173
Db 332 DFVSNIAATDSGAGVFTKENLSCTNTNSLQFLKN-----SAGQHGGG-AVVTQMTSVT 382
Qy 174 EARQRIELLEPKGTGE-----SKGAGESKGVGELRESNGAENTTETOTSTSTSSLR 226
Db 383 NTTSE--SITTPPLVGEVIFSENTAKHGCGGICTNKLSLN-----LKTVTLTKNSAKE 434
Qy 227 SDPKLWLALGTVATGLIGLAATGIVQALALTPEDSPPTTDP----- 268
Db 435 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFPAS 480
Qy 269 ---DAAASATETATRDQL--TKEAFQNPQDNQKVNIDELGNAIPSGVLKDDVVANIEEQA 323
Db 481 TAEPAAAPSLTE-AESDQTDQDETSDTNSDID-VSIENILN-----VAINQNTSAK 528
Qy 324 AA-----GEEAKQQAENNAQAQKYDEQAQKQBELKVSSGAGYGLSGLALILG-----GG 374
Db 529 KGGAIYGGKAKLSRINN-----LELSGNSQDVGGGLCLITESVEFDA 570
Qy 375 IGVAVT-----AALHRKNQVPEQTTTTTTTTTTTTSARTVEN---KPANNTPAQG-- 420
Db 571 IGSLLSHVNSAAKEGGVHISKTVLNLKSTFTPADNTVKAIVESTPEAPEIIPVEGEE 630
Qy 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459
Db 631 STATENPNSNTEGSSANTNLEGSGQDTADTGTGVVNNESQDTSDTGNAESGEQLQDSTQS 690
Qy 460 DVKTSLHDSQVPTNSNTSVQNMGTDSVVTSTIQHPPRDTT---DNGARLLGNPSAGI 515
Db 691 NEENTLPNSSIDQSNENTDESSDSHTEITDESVSSSKSGSSTPDQDGAASSGAPS-GD 749
Qy 516 QSTYAR--LALSGGLRHDMGGLTGGSNSAVNTSNPPAPGS 554
Db 750 QSISANACLAKSAASTDSSPVSNSGSDVDTASSDNDPSSS 790

RESULT 8
US-09-598-419-195
; Sequence 195, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 195
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-195

Query Match
  5.8%; Score 163.5; DB 4; Length 821;
Best Local Similarity 21.9%; Pred. No. 0.00035;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;

Qy 79 SEITLNDGFEVLHDGPLDNLNRQIGSSVFRVETQED--GK----- 117
Db 276 TEQTKSNG-----NQDGSSETKDTQVSESPESPDDVLGKGGGIYTEKSLTITGITGTI 331
Qy 118 ----HIAVQQRNGVETSVVLSDQEVARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173
Db 332 DFVSNIAATDSGAGVFTKENLSCTNTNSLQFLKN-----SAGQHGGG-AVVTQMTSVT 382
Qy 174 EARQRIELLEPKGTGE-----SKGAGESKGVGELRESNGAENTTETOTSTSTSSLR 226
```

Db 383 NTTSE--SITTPPLVGEVIFSENTAKHGCGICTNKLSLN-----LKTVTLTKNSAKE 434  
QY 227 SPKMLWALGTWATGLIGLAATGIVQALALTEPSPPTTDP----- 268  
Db 435 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFFAS 480  
QY 269 ---DAAASATETATRDQL--TKEAFONPDNOKVNDDELGNALIPSGVLKDDVVANIEEOKAK 323  
Db 481 TAEPAPASLITE--AESDQTDQETSDTNSDID--VSIENILN-----VAINQNTSAK 528  
QY 324 AA-----GEEAKQAIENNAQAQKDYDEOQAKRQBELKVSSGAGYGLSGALILG-----GG 374  
Db 529 KGGAIYKKAKLSRINN-----LELSGSSQDVGGGLCLTESVEFDA 570  
QY 375 IGVAVT-----AALHRKNQPVQETTTTTTTTTTSARTVEN---KPAANTPAQG-- 420  
Db 571 IGSLLSHYNSAAKEGGVIHKSIVTLNLSKSTFTPADNTVKAIVESTPEAPEIIPVVEGEE 630  
QY 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459  
Db 631 STATENPNSNTEGSSANTNLESGQDGTADTGTGVVNNESQDSDTGNAESGLOQDSTQS 690  
QY 460 DVKTSLHDSQVPTNSNTSVQNMGNMTDSVYVSTIOHPPRDIT---DNGARLLGNPSAGI 515  
Db 691 NEENTLPSNSIDQSNENTDESSDHSHEITDESVSSSKSGSTPDQCGAASSGAPS--GD 749  
QY 516 QSTYAR--LALSGGLRHDMGGLTGGNSNAVNTSNPPAPGS 554  
Db 750 QSISANACLAKEYAASTDSPPVNSSGSDVTASSNDPSSS 790

## RESULT 9

US-09-556-877-179  
; Sequence 179, Application US/09556877  
; Patent No. 6432916  
; GENERAL INFORMATION:  
; APPLICANT: Probat, Peter  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yaser  
; APPLICANT: Fling, Steve  
; APPLICANT: Maisonneuve, Jeff  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469CS  
; CURRENT APPLICATION NUMBER: US/09/556,877  
; CURRENT FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 305  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 179  
; LENGTH: 1776  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-09-556-877-179

Query Match 5.8%; Score 163.5; DB+4; Length 1776;  
Best Local Similarity 21.9%; Pred. No. 0.0011;  
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;  
QY 79 SEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETOED--GK----- 117  
Db 289 TEQTKSNG-----NQDGSSETKDTQVSESPESPSTPDDVLGKGGIYTEKSLTITGITGTI 344  
QY 118 ---HIAVQRNGVETSVVLSQDEYARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173  
Db 345 DFVSNIAATDSGAGVFTKENLSCNTNLSQFLKN-----SAGQHGGG--AYVTQMSVT 395  
QY 174 EARQRIELLEPKGTGE-----SKGAGESKGVGELRESNGBNTTETOTSTSTSLR 226  
Db 396 NTTSE--SITTPPLVGEVIFSENTAKHGCGICTNKLSLN-----LKTVTLTKNSAKE 447  
QY 227 SPKMLWALGTWATGLIGLAATGIVQALALTEPSPPTTDP----- 268

Db 448 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFFAS 493  
QY 269 ---DAAASATETATRDQL--TKEAFONPDNOKVNDDELGNALIPSGVLKDDVVANIEEOKAK 323  
Db 494 TAEPAPASLITE--AESDQTDQETSDTNSDID--VSIENILN-----VAINQNTSAK 541  
QY 324 AA-----GEEAKQAIENNAQAQKDYDEOQAKRQBELKVSSGAGYGLSGALILG-----GG 374  
Db 542 KGGAIYKKAKLSRINN-----LELSGSSQDVGGGLCLTESVEFDA 583  
QY 375 IGVAVT-----AALHRKNQPVQETTTTTTTTTTSARTVEN---KPAANTPAQG-- 420  
Db 584 IGSLLSHYNSAAKEGGVIHKSIVTLNLSKSTFTPADNTVKAIVESTPEAPEIIPVVEGEE 643  
QY 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459  
Db 644 STATENPNSNTEGSSANTNLESGQDGTADTGTGVVNNESQDSDTGNAESGLOQDSTQS 703  
QY 460 DVKTSLHDSQVPTNSNTSVQNMGNMTDSVYVSTIOHPPRDIT---DNGARLLGNPSAGI 515  
Db 704 NEENTLPSNSIDQSNENTDESSDHSHEITDESVSSSKSGSTPDQCGAASSGAPS--GD 762  
QY 516 QSTYAR--LALSGGLRHDMGGLTGGNSNAVNTSNPPAPGS 554  
Db 763 QSISANACLAKEYAASTDSPPVNSSGSDVTASSNDPSSS 803

## RESULT 10

US-09-620-412C-179  
; Sequence 179, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 179  
; LENGTH: 1776  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-09-620-412C-179

Query Match 5.8%; Score 163.5; DB 4; Length 1776;  
Best Local Similarity 21.9%; Pred. No. 0.0011;  
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;  
QY 79 SEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETOED--GK----- 117  
Db 289 TEQTKSNG-----NQDGSSETKDTQVSESPESPSTPDDVLGKGGIYTEKSLTITGITGTI 344  
QY 118 ---HIAVQRNGVETSVVLSQDEYARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173  
Db 345 DFVSNIAATDSGAGVFTKENLSCNTNLSQFLKN-----SAGQHGGG--AYVTQMSVT 395  
QY 174 EARQRIELLEPKGTGE-----SKGAGESKGVGELRESNGBNTTETOTSTSTSLR 226  
Db 396 NTTSE--SITTPPLVGEVIFSENTAKHGCGICTNKLSLN-----LKTVTLTKNSAKE 447  
QY 227 SPKMLWALGTWATGLIGLAATGIVQALALTEPSPPTTDP----- 268  
Db 448 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFFAS 493  
QY 269 ---DAAASATETATRDQL--TKEAFONPDNOKVNDDELGNALIPSGVLKDDVVANIEEOKAK 323  
Db 494 TAEPAPASLITE--AESDQTDQETSDTNSDID--VSIENILN-----VAINQNTSAK 541  
QY 324 AA-----GEEAKQAIENNAQAQKDYDEOQAKRQBELKVSSGAGYGLSGALILG-----GG 374  
Db 542 KGGAIYKKAKLSRINN-----LELSGSSQDVGGGLCLTESVEFDA 583

```
QY 375 IGAVT-----AALHRKNQPVQEQTTTTTTTTTTTTSARTVEN---KPANNTPAQG-- 420
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 584 IGSLLSHYNSAAKEGGVIHSTKVTILSNLKSFTTFADNTVKAIVESTPEAPEIIPVEGEE 643
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSSIGTVQ-----NPIYA 459
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 644 STATENPNSNTEGSSANTNLEGSGDGTADTGTGVVNNESQDTSDTGNAESGEQLQDSTQS 703
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 460 DVKTSLHDSQVPTNSNTSVQNMGNMNTDSVVYSTIQHPRPDRTT---DNGARLLGNPFSAGI 515
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 704 NEENTLPNSSIDQSNENTDESSDHTBEITDESVSSSKSGSSTPDQCGAASSGAPS-GD 762
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 516 QSYAR--LALSGGLRHDMGGLTGSNSAVNTSNPPAPGS 554
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 763 QSISANACLAASYAASTDSSPVNSSGSDVTASSDNDPSSS 803
|| : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 11
US-09-598-419-179
; Sequence 179, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 179
; LENGTH: 1776
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-179
Query Match 5.8%; Score 163.5; DB 4; Length 1776;
Best Local Similarity 21.9%; Pred. No. 0.0011;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;
QY 79 SEITUNDGEVULHDPGLDTLNRQGTSSVFRVETOED--GK----- 117
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 TEQTKSNG---NQDGSSETKDQTVSESPESPDPDLVGKGGIYTEKSLTITGITGTI 344
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 ---HIAVQQRNGVETSVVLSDOEYARLQSIDPECKDFVFTGGRGGAGHAMVTVASDIT 173
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 DFVSNIAIDSGAGVFTKENLSCTNNSLQFLKN-----SAGOHGGG-AYVTQTMGVT 395
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 EARQRILELLEPKGTGE-----SKGAGESKGVGELRESNSGAENTTETOTSTSTSLR 226
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 396 NTTSE--SITTPPLVGEVIFSENTAKHGGGICTNKLISN-----LKTVTLTKNSAKE 447
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 227 SPKMLWALGTVATGLIGLAATGIVQALALTPEDPSPTTDP----- 268
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 448 SG-----GAFTDLASIPTD-----TPESSTPGSSSPASPPEVVASAKINRFPAS 493
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 269 ---DAAASATEATRDQL--TKEAFQNPQNKVNTDELGNALPSGVLKDDVNVANIEEQA 323
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 494 TAEPAAPSLTE-AESDQTDQTTSDTNSDID-VSIENILN-----VAINQNTSAK 541
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 AA---GGEAKQQAENNAQAQKYDEQQAQRQEBLKVSSGAGYGLSGALILG-----GG 374
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 542 KGAIVGKAKULSRNN-----LELSGNSQDVGGGLCTESVEFDA 583
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 375 IGAVT-----AALHRKNQPVQEQTTTTTTTTTTTTSARTVEN---KPANNTPAQG-- 420
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 584 IGSLLSHYNSAAKEGGVIHSTKVTILSNLKSFTTFADNTVKAIVESTPEAPEIIPVEGEE 643
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSSIGTVQ-----NPIYA 459
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 644 STATENPNSNTEGSSANTNLEGSGDGTADTGTGVVNNESQDTSDTGNAESGEQLQDSTQS 703
|| : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 460 DVKTSLHDSQVPTNSNTSVQNMGNMNTDSVVYSTIQHPRPDRTT---DNGARLLGNPFSAGI 515
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 704 NEENTLPNSSIDQSNENTDESSDHTBEITDESVSSSKSGSSTPDQCGAASSGAPS-GD 762
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 516 QSYAR--LALSGGLRHDMGGLTGSNSAVNTSNPPAPGS 554
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 763 QSISANACLAASYAASTDSSPVNSSGSDVTASSDNDPSSS 803
|| : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 12
US-09-248-796A-15936
; Sequence 15936, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15936
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (322),(418)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-15936
Query Match 5.6%; Score 159.5; DB 4; Length 441;
Best Local Similarity 21.9%; Pred. No. 0.00029;
Matches 90; Conservative 48; Mismatches 162; Indels 111; Gaps 15;
QY 183 LEPKGTGESKAGSGKGVGELRESNSGAENTTETOTSTSSLSRSDPKLWALGTVATGL 242
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 88 LEKENTENNQTDNESAGVDEDNSTNKFDDIDAKSELSTSSRLRN-----GK 134
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 IGLAATGIVQALALTPEDPSPTT-----TDPDAAASATETA---TRDQLTKEAFQNP 292
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 FSNLTLDLTKEMKKKSPSLNTENSLSVNNTNTASLSASNTLNKTPSKLDTASQSPN 194
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 NQKNIDELGNAIPSGVLKDDVANI BEQAKAAGEAKQQAENNAQAQKYDEQQAQRQ 352
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 TNSTLVSTFKKIGKWYFKDD-----GSSKKQHDEKLGYP 230
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 EELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQP-VEQTTTT---TTTTTTSART 407
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 PEMISSRRAPKSSSMTGLGSDSTRYTLATENKDLPIESANTTSASVNTATTAPSL 290
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 408 VENKANNTPAQGNVDTPGSEDPTMESRRSSMASTSTFFDTSIGTQVONPYADVK---TS 464
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 LSNK-GITTTAPVNVISLSNNFTVPSVAVGADXTT-----TAPAPVANSKLTSTP 339
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 LHDSQVPTNSNTSVQNMGNMNTDSVVYSTIQH-----PPRDTDN-----GAR 506
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 L-DNQTSPSYNFRSLDNN--NEDLSIFSVAYQNAAQIIPSLTPPLPPRPDI SDNLSIRSGS 396
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 507 LLGNPSAGIQSYARLALSGGLRHDMGGLTGSNSAVNT---SNNPPAPGS 554
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 397 LVGEPFTLSLQ-----PNSPIKTKLXRPKPPPPPKS 426
|| : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 13
US-09-248-796A-22338
; Sequence 22338, Application US/09248796A
; Patent No. 6747137
```

```

; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22338
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-22338

Query Match          5.6%; Score 158; DB 4; Length 556;
Best Local Similarity 21.8%; Pred. No. 0.00054;
Matches 93; Conservative 44; Mismatches 200; Indels 90; Gaps 13;

QY 169 ASDITRQRILLELPKGTGSGKAGE-----SKGVGELRESNGAENT-----213
DB 71 STDITSASS--TDEQSSSGTGQSSTEDPIDSTESDTSSATSDSSATNTDNTNSES 128
QY 214 TETQSTSTSSLSRSDPKLMLALGTGATGLGLAATGIVQA-----LALTP 258
DB 129 TDSSTATDSSSTDSNTASSTETNTDVTSDTSNTGATESSTATDNTDATDSSTVSETG 188
QY 259 EPDPTTTPDAAA-----SATETATRDQLTKEAFQNPNDKVNIDELGNAI 305
DB 189 ATDSSSTATDNTGATSSSDTSNTGATESSTATDNTSANTDNTGTSNTATNTD--NTAT 247
QY 306 PGVLKDDVVVANIEQAKAAGEAKQQAJENNAQAOKKYDEQAKRQBELKVSAGYGL 365
DB 248 DTSSSTETATNTD-----GTETNTGATETNTDTSASNTDDNT-----GS 287
QY 366 SGALLIGGGIGVAVTAALHRKQNPVQTTTTTTTTTTSARTVENKPNANTPAQGNVDP 425
DB 288 NTATNTGG-----TDNTDNTGTDNTGTDNTGTDNTGTDNTGTDNTGTDNTG 329
QY 426 GSEDVTMESRRSSMASTSTFFDTSIGTVQNPY--ADVKTSLHDSQVPTSNSTSVQNGW 483
DB 330 NTGATETATNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNT 389
QY 484 NTDSVVYSTIQHPPRDT--TDNGARLLGNPSAGIQSTYARLALSG--GLRHDMGGLTGGSNS 541
DB 390 NTDNV--STTKDIPSPSTINEGSGNGSGNGSGNGSGNGSGNGSGNGSGNGSGNGSG 447
QY 542 AVNTSN 548
DB 448 SCNGSDN 454

RESULT 14
US-09-248-796A-17307
; Sequence 17307, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17307
; LENGTH: 827

```

```

; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-17307

Query Match          5.5%; Score 157.5; DB 4; Length 827;
Best Local Similarity 22.5%; Pred. No. 0.0011;
Matches 126; Conservative 63; Mismatches 251; Indels 119; Gaps 24;

QY 56 ADSGDNRASDVPLPVNMRLLAASITLNDGFEVLH--DHGPLDITLNRQIGSVFRVETQ 113
DB 37 ADNIDKRA---GAIGNFFRDFTNISFGNDNLEVNQPNSTNGATST--GHFFGSPSTSTH 91
QY 114 EDGKHIAVGQNGVETSVVLSQOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDIT 173
DB 92 -----QQTGFTSNVNTKSSQONS--PS-----TSPSTVAAAATSSPVA 133
QY 174 EARQRILELLEPKGTGSGKAGESKGVGELRESNGAENTTTQTSTSTSSLSRSDPKLML 233
DB 134 STR-----PASTSEQKQEESTA---RQSTSPATATTSTNTPPSSTSKETPTSTNT 181
QY 234 ALGTVA-----TGLIGLAATGIV-----QALALTP-----EPDPTTTPDAAA 272
DB 182 AQTSSANNQOSSNTAAPTSTVIQPSSTSEVHVQSQOTSTTNTPTTSSPNTPTTSEAAPTT 241
QY 273 SATETATRDQLTKEAFQNPNDKV-----NIDELGNA-----IPSGVLKDDVVANIEQ 322
DB 242 SAAPT-----TSEAPVPTSTSEVVPNTPTTSEAPNTPTTSEAPVPTSTSEVVPNTPTTS 295
QY 323 KAAGBEAKQ--AIENNAQAOKKYDEQQA---KQBELKVSAGYGLSGALILGGGIG 376
DB 296 KAPNTPTTSEAPATTTSEAPNTPTTSEAPVPTTSEVVPNTTSTQGDVSTS-----S 348
QY 377 VAVTAALHRKQNPVQTTTTTTTTTTSARTVENKPN---NTPA-----QGNVD 423
DB 349 TSVTEQTTLTSLTLLPPTTALTQTSTPEASDPKPSSTSIETPSTSTFEQDPTTTSSVG 408
QY 424 TPGSDDVTMESRRSSMASTSTFFDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQNGW 483
DB 409 TPSSPOPOPTTSELAVTSNS--PTQESTSLVEP---TTSSLESNTPTNPTSTSEAOQS 463
QY 484 NTDSVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGL-----RHDMGGLTG 537
DB 464 TSAS-----QAPPTTSSAPAPELSSNADPSNLVHSSETTSLVNPSTDQIDSSSTTD 517
QY 538 GNSAVN--TSNPPAPGS 554
DB 518 AVSQATTEPTSENTPTAAS 536

RESULT 15
US-09-949-016-10491
; Sequence 10491, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10491
; LENGTH: 2045
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-10491

```

```

Query Match      5.4%; Score 152.5; DB 4; Length 2045;
Best Local Similarity 18.6%; Pred. No. 0.0099;
Matches 134; Conservative 87; Mismatches 295; Indels 205; Gaps 23;

QY 3  IGNLGHNPVNNISIPAPPLPSOTDAGGRG---QLINSTGPLGSRALFPPVRNSMADSG 59
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
687  ISNLG---KMSVWQTKPVQTSVAVTQASTGPVTQIIQTKGPLPAGTIL-----KLVTSA 738
QY 60  DNRASDV-----PGLPVNPMELAASEI---TLNDGFEVLHDHGPDLTLNRQIGSSVFRV 110
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
739  DGKPTTIITTOASGAGTKPTILGISSVSPSTPKPFTTIKTIIPMSAIIQAGAT---- 794
QY 111  ETQEDGKHAVGQRMGVETSVVLSDOEY-----ARLOSIDP-----EGKDKFVF 154
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
795  -----GVTSSPGIKSPITIIITKMTSGTGAPAKIITAVPKIATGHGQGGVTQVVL 845
QY 155  TGRRGAGHAMVTV-----ASDIT-EARQRIELLEPKGTGSGKAGESKGVGELRES 206
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
846  KGAPQPGTILRTVPMGGVRLVTPVTVSAVKPAVTLVVKGTGTGVTGTVSTSLA 905
QY 207  NSGAENTTETQSTSTSSLRSDPKLWALGTVAT-----GLIGLAATGIV 251
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
906  GAGHSTASLATPITT-----LGTIATLSSQVINPITVSAAGTTLTAAGGL 954
QY 252  QALALTPEP-----DSPTTDPDAAAASATETATRDQ 282
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
955  TPTTITMQPVSOPTQVTLITAPSGVEAQPVHDLVPSILASPTTEQPTATVTIADSGQGDV 1014
QY 283  LTKEAFQNPNDQKVINDELGNAIPSGVLKDDVANI-----EEQAKAAGBEA 329
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
1015  QPGTVTLVCSNPPCETHETGT---TNTATTTVANLGGHPQPTQVFCDRQEAASLVT 1071
QY 330  KQOAIENNA-----QAQKYDEQAKRQEELKVSSGAGYGLSGALIL 371
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
1072  STVGQONGSVHGCNPPCETHETGTNTATTAMSSVGANHQDARRACAAG---TPAVIR 1128
QY 372  GGGIGVAVTAALHRKNQPVQEQTTTTTTTTTSAR-----TVENKPANNTPA- 418
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
1129  ISVATGALENAAGSKSQCOCTROTTSATSTMTVMATGAPCSAGPLGPSNAREPGGRSPAF 1188
QY 419  -----QGNVDTPGSEDTMESRRSSMASTSTFFDTSIGTVQNPYADVKTSLHDSQV 470
Db    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
1189  VQLAPLSKVKRLSSPSIKDLPAGRHSHAVSTAA--MTRSSVGAGEPRMAPVCESLQGGSP 1246
QY 471  PTSNNTSVQNMGNDSVYVYSTIOHPPRDTTNGA-----RLGNP-----S 512
Db    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
1247  STTIVTVALEALLCPSATVTQVCNPPCETHETGTNTATTATTSNAGSAQRCVCSNPPCETHE 1306
QY 513  AGIQSTYARLALSGLRHDMGG-----LTGGSNSAVNTSNPPAPGSHRF 557
Db    |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
1307  TGTTHATATATNGGTGQPEGQPPAGPCETHQTSTGTMTSVSGALLPDATSSHRT 1366
QY 558  V 558
Db    |
1367  V 1367

```

Search completed: May 13, 2005, 11:49:11  
Job time : 46 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:16:56 ; Search time 178 seconds  
(without alignments)  
1605.282 Million cell updates/sec

Title: US-09-189-415b-11

Perfect score: 2840

Sequence: 1 MPIGNLGHNPVNSIPPAP.....SNSAVNTSNPPAGSHRFV 558

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2840	100.0	558	2	Q7DB77	Q7db77 escherichia
2	2840	100.0	558	2	Q9R396	Q9r396 escherichia
3	2824	99.4	558	2	O85506	O85506 escherichia
4	1843	64.9	538	2	Q47014	Q47014 escherichia
5	1835	64.6	538	2	Q47016	Q47016 escherichia
6	1824	64.2	538	2	O85508	O85508 escherichia
7	1613.5	56.8	551	2	O68258	O68258 escherichia
8	1594	56.1	552	2	Q9KWH9	Q9kwh9 escherichia
9	1585	55.8	550	2	O52147	O52147 escherichia
10	1569.5	55.3	547	2	Q7BHL5	Q7bhl5 escherichia
11	1569.5	55.3	547	2	Q9WKK1	Q9wxk1 escherichia
12	1568.5	55.2	547	2	O9ETI1	O9etil citrobacter
13	1531.5	53.9	549	2	O50190	O50190 escherichia
14	1146.5	40.4	367	2	Q79C12	Q79c12 escherichia
15	215.5	7.6	1374	2	Q7YU77	Q7yu77 drosophila
16	214.5	7.6	1323	2	Q7KXZ0	Q7kxz0 drosophila
17	214.5	7.6	1376	2	Q9VI63	Q9vi63 drosophila
18	208.5	7.3	1321	2	O963D1	O963d1 drosophila
19	207.5	7.3	1015	2	O8SZW9	O8szw9 drosophila
20	200	7.0	2332	2	Q8LFX6	Q8lfx6 caenorhabdi
21	196	6.9	2338	2	Q759S3	Q759s3 ashbya goss
22	195.5	6.9	1254	2	Q94185	Q94185 caenorhabdi
23	194.5	6.8	3295	2	O66GT3	O66gt3 rattus norv
24	187.5	6.6	3550	2	O66GT4	O66gt4 rattus norv
25	187	6.6	1291	2	Q77261	Q77261 drosophila
26	187	6.6	1300	2	Q9W5E0	Q9w5e0 drosophila
27	185	6.5	2310	2	Q8CMU7	Q8cmu7 staphylococ
28	184.5	6.5	3135	2	Q7KHU2	Q7khu2 drosophila
29	184	6.5	1961	2	Q7PQ78	Q7pq78 anopheles g
30	182	6.4	574	2	Q7R861	Q7r861 neurospora
31	182	6.4	3080	2	Q9V602	Q9v602 drosophila

32	182	6.4	3109	2	Q9BMO0	Q9bmq0 drosophila
33	180	6.3	596	2	Q6UXC5	Q6uxc5 homo sapien
34	178.5	6.3	3443	2	Q6UJZM8	Q6jzm8 mus musculu
35	177	6.2	1351	2	Q6GDE9	Q6gde9 staphylococ
36	176	6.2	2283	2	Q8VQ99	Q8vg99 staphylococ
37	174.5	6.1	1241	2	Q75JCO	Q75jco dictyosteli
38	174.5	6.1	2634	2	O63KC6	O63kc6 burkholderi
39	174	6.1	1176	2	Q869T7	Q869t7 dictyosteli
40	174	6.1	1535	2	Q62D27	Q62d27 burkholderi
41	174	6.1	2275	2	Q8NUJ3	Q8nuj3 staphylococ
42	174	6.1	2275	2	Q6G620	Q6g620 staphylococ
43	174	6.1	5899	2	Q74GG7	Q74gg7 geobacter s
44	173	6.1	1376	2	Q7SSH8	Q7ssh8 neurospora
45	172.5	6.1	2271	2	Q99QY4	Q99qy4 staphylococ

## ALIGNMENTS

### RESULT 1

Q7DB77	PRELIMINARY;	PRT;	558 AA.
AC	Q7DB77; Q7A9Q1;		
DT	05-JUL-2004 (Tremblrel. 27, Created)		
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)		
DT	25-OCT-2004 (Tremblrel. 28, Last annotation update)		
DE	Putative translocated intimin receptor protein (Translocated intimin receptor Tir).		
GN	Name=tir; OrderedLocusNames=ECs4561, Z5112;		
OS	Escherichia coli O157:H7.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Escherichia.		
OX	NCBI_TaxID=83334;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;		
RX	MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;		
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,		
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,		
RA	Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,		
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,		
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,		
RA	Welch R.A., Blattner F.R.;		
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.;"		
RL	Nature 409:529-533(2001).		
RN	[2]		
RC	SEQUENCE FROM N.A.		
RX	STRAIN=O157:H7 / RIMD 050952 / EHEC;		
RA	MEDLINE=21156231; PubMed=11258796;		
RA	Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,		
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,		
RA	Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,		
RA	Kuhara S., Shiba T., Hattori M., Shinagawa H.;		
RT	"Complete genome sequence of enterohaemorrhagic Escherichia coli		
RT	O157:H7 and genomic comparison with a laboratory strain K-12.;"		
RL	DNA Res. 8:11-22(2001).		
DR	EMBL; AE005595; AAG58825.1; --		
DR	EMBL; AP002566; BAB37984.1; --		
DR	GO; GO:0004872; F:receptor activity; IEA.		
DR	GO; GO:0007155; P:cell adhesion; IEA.		
DR	InterPro; IPR003536; Tir_receptor.		
DR	Pfam; PF07489; Tir_receptor C; 1.		
DR	Pfam; PF03549; Tir_receptor M; 1.		
DR	Pfam; PF07490; Tir_receptor N; 1.		
DR	PRINTS; PR01370; TRANSINTINR.		
SQ	SEQUENCE 558 AA; 58022 MW; 99C417222D4B4AAL CRC64;		

Query Match 100.0%; Score 2840; DB 2; Length 558;

Best Local Similarity 100.0%; Pred. No. 3e-142;

Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGHNPVNSIPPAPLPSOTDAGRGQLNSTGPIGSRALTPVNSVADSGD 60

```
Db 1 MPIGNLGNPNVNNISIPAPPSPQSDGAGRGQLINSTGPGSRALFTPVVNSMADSGD 60
Qy 61 NRASDVPLGPNVPMRLAAASEITLNDGFEVLHDGPDLTNROIGSSVRFVETQEDGKHIA 120
Db 61 NRASDVPLGPNVPMRLAAASEITLNDGFEVLHDGPDLTNROIGSSVRFVETQEDGKHIA 120
Qy 121 VQORNGVETSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQIL 180
Db 121 VQORNGVETSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQIL 180
Qy 181 ELLEPKGTGESKGAGESKGVGELRESNCSAENTTETQTSSTSLRSDPKLWALGTAT 240
Db 181 ELLEPKGTGESKGAGESKGVGELRESNCSAENTTETQTSSTSLRSDPKLWALGTAT 240
Qy 241 GLIGLAATGIVQALALTPEDSPPTTDPDAASATETATRDQLTKEAFQNDQKNVDE 300
Db 241 GLIGLAATGIVQALALTPEDSPPTTDPDAASATETATRDQLTKEAFQNDQKNVDE 300
Qy 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAQIENNAQAQKQYDEQAQKQEEELKVSSG 360
Db 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAQIENNAQAQKQYDEQAQKQEEELKVSSG 360
Qy 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPANNTPAQG 420
Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPANNTPAQG 420
Qy 421 NVDTFGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db 421 NVDTFGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Qy 541 SAVNTSNPPAPGSHRFV 558
Db 541 SAVNTSNPPAPGSHRFV 558

RESULT 2
Q9R396 PRELIMINARY; PRT; 558 AA.
AC Q9R396;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Translocated intimin receptor Tir (L0027).
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86/24;
RX MEDLINE=99242825; PubMed=10225900;
RA Devlinney B., Stein M., Reineheid D., Abe A., Ruschkowski S.,
RA Finlay B.B., Blattner F.R.;
RT "Enterohemorrhagic Escherichia coli O157:H7 produces Tir, which is
RT translocated to the host cell membrane but is not tyrosine
RT phosphorylated.";
RL Infect. Immun. 67:2389-2398(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC43895, and EDL933;
RX MEDLINE=98339885; PubMed=9673266;
RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
RA Kaper J.B., Blattner F.R.;
RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT Escherichia coli O157:H7";
RL Infect. Immun. 66:3810-3817(1998).
DR EMBL; AF125993; AAD29391.1; -.
DR EMBL; AF071034; AAC31506.1; -.

```

```
PIR; A98199; A98199.
PIR; E86045; E86045.
DR HSSP; Q9KWH9; 1P02.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor C; 1.
DR Pfam; PF03549; Tir_receptor M; 1.
DR Pfam; PF07490; Tir_receptor N; 1.
DR PRINTS; PR01370; TRSINTIMNR.
KW Receptor.
SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4B4AA1 CRC64;

Query Match 100.0%; Score 2840; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 3e-142;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPIGNLGNPNVNNISIPAPPSPQSDGAGRGQLINSTGPGSRALFTPVVNSMADSGD 60
Db 1 MPIGNLGNPNVNNISIPAPPSPQSDGAGRGQLINSTGPGSRALFTPVVNSMADSGD 60
Qy 61 NRASDVPLGPNVPMRLAAASEITLNDGFEVLHDGPDLTNROIGSSVRFVETQEDGKHIA 120
Db 61 NRASDVPLGPNVPMRLAAASEITLNDGFEVLHDGPDLTNROIGSSVRFVETQEDGKHIA 120
Qy 121 VQORNGVETSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQIL 180
Db 121 VQORNGVETSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQIL 180
Qy 181 ELLEPKGTGESKGAGESKGVGELRESNCSAENTTETQTSSTSLRSDPKLWALGTAT 240
Db 181 ELLEPKGTGESKGAGESKGVGELRESNCSAENTTETQTSSTSLRSDPKLWALGTAT 240
Qy 241 GLIGLAATGIVQALALTPEDSPPTTDPDAASATETATRDQLTKEAFQNDQKNVDE 300
Db 241 GLIGLAATGIVQALALTPEDSPPTTDPDAASATETATRDQLTKEAFQNDQKNVDE 300
Qy 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAQIENNAQAQKQYDEQAQKQEEELKVSSG 360
Db 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAQIENNAQAQKQYDEQAQKQEEELKVSSG 360
Qy 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPANNTPAQG 420
Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPANNTPAQG 420
Qy 421 NVDTFGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db 421 NVDTFGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Qy 481 NMGNNTDSVVYSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDWGGLTGGSN 540
Db 481 NMGNNTDSVVYSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDWGGLTGGSN 540
Qy 541 SAVNTSNPPAPGSHRFV 558
Db 541 SAVNTSNPPAPGSHRFV 558

RESULT 3
O85506 PRELIMINARY; PRT; 558 AA.
AC O85506;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DE 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Translocated intimin receptor Tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=95SF2;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (tir) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586 (1998).
DR EMBL; AF070067; AAC69314.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMINR.
KW Receptor.
SQ
SEQUENCE 558 AA; 58175 MW; CA2CDDAC94527C2E CRC64;
Query Match 99.4%; Score 2824; DB 2; Length 558;
Best Local Similarity 99.5%; Pred. No. 2.1e-141;
Matches 555; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MPIGNLGHNPVNNNSIPPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSG 60
DB 1 MPIGNLGHNPVNNNSIPPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSG 60
QY 61 NRASDVPGLPVNPMLRAASEITLNDGFVLHDHGPLDTLNRIQIGSSVFRVTOEDGKHIA 120
DB 61 NRASDVPGLPVNPMLRAASEITLNDGFVLHDHGPLDTLNRIQIGSSVFRVTOEDGKHIA 120
QY 121 VQORNGVETSVLVSDQYARLOSIDPEGKDKVFTGGRGAGHAMVTVASDITEARQRL 180
DB 121 VQORNGVETSVLVSDQYARLOSIDPEGKDKVFTGGRGAGHAMVTVASDITEARQRL 180
QY 181 ELLEPKGTGSKGAGSKGVGELRENSGAENTTETQSTSSLSRSDPKLWALGTAT 240
DB 181 ELLEPKGTGSKGAGSKGVGELRENSGAENTTETQSTSSLSRSDPKLWALGTAT 240
QY 241 GLIGLAATGIVQALALTPEDSPFTTDPDAASATETATRDQITKEAFQNPQKNIDE 300
DB 241 GLIGLAATGIVQALALTPEDSPFTTDPDAASATETATRDQITKEAFQNPQKNIDE 300
QY 301 LGNAIPSGVLKDDVANIIEQAKAAGEAKQAIIENNAQAQKYDEQQAQKEELKVSSG 360
DB 301 LGNAIPSGVLKDDVANIIEQAKAAGEAKQAIIENNAQAQKYDEQQAQKEELKVSSG 360
QY 361 AGVGLSGALILGGIGVAVTAALHRKNQVPEQTTTTTTTTTTTTTSARTVENKPNNTPAQ 420
DB 361 AGVGLSGALILGGIGVAVTAALHRKNQVPEQTTTTTTTTTTTTTSARTVENKPNNTPAQ 420
QY 421 NVDTPGSEDTMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
DB 421 NVDTPGSEDTMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
QY 481 NMGTDSVYVYTIQHPRDTTNGARLLGNPAGIOSTYARLALSGGLRHDMGGLTGGSN 540
DB 481 NMGTDSVYVYTIQHPRDTTNGARLLGNPAGIOSTYARLALSGGLRHDMGGLTGGSN 540
QY 541 SAVNTSNPPAGSHRFV 558
DB 541 SAVNTSNPPAGSHRFV 558
RESULT 4
ID Q47014 PRELIMINARY; PRT; 538 AA.
AC Q47014;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Translocated intimin receptor tir (Translocated intimin co-receptor)

```

```

DE GN Name=tir; Synonyms=espE;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC 84/110/1, and E65/56;
RA Krejany E.O.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=95ZG1;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (tir) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=B10;
RX MEDLINE=20187493; PubMed=10722617;
RI DOI=10.1128/IAI.68.4.2171-2182.2000;
RA Marches O., Nougayrede J.P., Boullier S., Mainil J., Charlier G.,
RA Raymond I., Pohl P., Boury M., De Rycke J., Milton A., Oswald E.;
RT "Role of tir and intimin in the virulence of rabbit enteropathogenic
RT Escherichia coli serotype O103:H2.";
RL Infect. Immun. 68:2171-2182 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=413/89-1;
RX MEDLINE=982994040; PubMed=9632251;
RA Deibel C., Kraemer S., Chakraborty T., Ebel F.;
RT "EspE, a novel secreted protein of attaching and effacing bacteria, is
RT directly translocated into infected host cells where it appears as a
RT tyrosine-phosphorylated 90 kDa protein.";
RL Mol. Microbiol. 28:463-474 (1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=413/89-1;
RA Benkel P., Chakraborty T.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U59502; AAC32028.2; -.
DR EMBL; AF070068; AAC69316.1; -.
DR EMBL; AF113597; AAF03080.1; -.
DR EMBL; AF132728; AAD27868.1; -.
DR EMBL; AJ223063; CAA11065.1; -.
DR EMBL; AJ277443; CAC81869.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMINR.
KW Receptor.
SQ
SEQUENCE 538 AA; 55420 MW; 31D7A8E227B3D06C CRC64;
Query Match 64.9%; Score 1843; DB 2; Length 538;
Best Local Similarity 65.9%; Pred. No. 1.5e-89;
Matches 371; Conservative 61; Mismatches 101; Indels 30; Gaps 6;
QY 1 MPIGNLGHNPVNNNSIPPAPPLPSQTDGAGG-RGQLINSTGLGSRALFTPVNSMADSG 59
DB 1 MPIGNLGHNPVNNNSIPPAPPLPSQTDGAGG-RGQLINSTGLGSRALFTPVNSMADSG 60
QY 60 DNRASDVPGLPVNPMLRAASEITLNDGFVLHDHGPLDTLNRIQIGSSVFRVTOEDGKH 119
DB 61 DSRASDI FGLPTNPLRFRASEVSLHGALEVLHDKGLDTLNSAIGSSLPFRVTRDGGSHV 120

```



```

QY 420 GNVDPGSDTMSRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSV 479
DB 404 GNTDTSGPESPASRRNSASLASNGSDTSSGTGVENYADV-----GHPNDLSLARI 456
QY 480 QNMGMNTDSVV-----YSTIOHPRPDITDNGARLLGNPSAGIQSTYARLALSGGLRHDGMGL 535
DB 457 SEEPYDEVAADPNYSVIOHFSGNSPVTG-RLVGTGQGIQSTYALLASSGGLRLGMGL 515
QY 536 TGGNSAVNTSNPPAPGSHRFV 558
DB 516 TGGESAVSTANAAPTGPARFV 538

RESULT 6
O85508 PRELIMINARY; PRT; 538 AA.
AC O85508;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Translocated intimin receptor Tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EPEC87A;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RA "Translocated intimin receptors (Tir) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
DR EMBL; AF070069; AAC69318.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor.
SQ SEQUENCE 538 AA; 55602 MW; 447052A0E3214D6D CRC64;

Query Match
Best Local Similarity 64.2%; Score 1824; DB 2; Length 538;
Matches 369; Conservative 60; Mismatches 106; Indels 26; Gaps 6;

QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGG-RGOLINSTGPGSRALFTPVNRSMADSG 59
DB 1 MPIGNLGNPNVRLIIPAPPLPSQTDGAGGARNQLINSNGPMGSRLLFTPIRNSVADA 60
QY 60 DNRASDVPGLVPNMRLLAASEITLNDGFVLHDHGLDPLTNKQIGSSVFRVETQDGHX 119
DB 61 DSRASDIPGLPTNPLRFAASEVSLHGALEVLHDHKGGLDPLTNSAIGSLFRVETRDGSHV 120
QY 120 AVQORNGVTSVVLSDQEARLOSIDPECKDFVETGGGGAGHAMVTVASDITEARORI 179
DB 121 ATGQRKNGLETTVVLSDQEFFFSIQSLDPECKKFKVFTGGGGAGHAMVTVASDIAEARORI 180
QY 180 LLELPKPGTGESKAGESKGVGELRESNSGAENTTETOTSTSLRSDPKLWALGTVA 239
DB 181 IDKLEPKOTKETKEP-----DPSNGEKKIIEHTSTSTFSLRADPKLWLSLGHIA 231
QY 240 TGLIGLAATGIVQALALTPEDPSPTTTPDAAASATETATRDQLTKEAFQNPNDKQNVID 299
DB 232 AGLIGMAATGIAQAVALTPEPDPTTTPDPTAATAEAATKDRLTQEAQFDPDKQKVID 291
QY 300 ELGNAIPSGVLKDDVVVAVIEEQAKAAGEAKQQAENNAQAOKKYDEQAQRKEELKVSS 359

```

```

DB 292 ENGNAIPSGELIDDVVAQIAEQAAQEAQEAIESNSQAQKYDEQAHKREQEWSLSS 351
QY 360 GAGYGLSALILGGIGVAVTAALHRKNQPVQETTTTTTTTTTTSARTVENKPNANTPAQ 419
DB 352 GVGYGISALILGGIGAGVTAALHRKNQPAEQITTTT-----YVDNQPTNNSAQ 403
QY 420 GNVDPGSDTMSRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDS--QVPTSNST 477
DB 404 GNTDTSGPESPASRRNSASLASNGSDTSSGTGVENYADVGMPRNDLSLARIPEPIYD 463
QY 478 SVQNMGMNTDSVVYSTIOHPRPDITDNGARLLGNPSAGIQSTYARLALSGGLRHDGMGLTG 537
DB 464 EVAADPN-----YSVIOHFSGNSPVTG-RLVGTGQGIQSTYALLASSGGLRLGMGLTG 517
QY 538 GSN SAVNTSNPPAPGSHRFV 558
DB 518 GGESAVSTANAAPTGPARFV 538

RESULT 7
O68258 PRELIMINARY; PRT; 551 AA.
AC O68258;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Translocated intimin receptor.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99003184; PubMed=9529069;
RA Voss E., Paton A.W., Manning P.A., Paton J.C.;
RA "Molecular analysis of Shiga toxicogenic Escherichia coli O111:H-
RT proteins which react with sera from patients with hemolytic-uremic
RT syndrome.";
RL Infect. Immun. 66:1467-1472(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RA "Translocated intimin receptors (Tir) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
DR EMBL; AF025311; AAC69249.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor.
SQ SEQUENCE 551 AA; 56975 MW; 68C95F76BF0F44CC CRC64;

Query Match
Best Local Similarity 56.8%; Score 1613.5; DB 2; Length 551;
Matches 346; Conservative 68; Mismatches 126; Indels 29; Gaps 13;

QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGCA--GGGOLINSTGPGSRALFTPVNRSMADS 58
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGASGGAGQLINSTGALGSRLLFSLRNSADS 60
QY 59 GDNRASDVPGLVPNMRLLAASEITLNDGFVLHDHGLDPLTNKQIGSSVFRVETQDGH 116
DB 61 VDSR--DIPGLVPNPSRLATATSEICLLGGFVLHDHKGGLDPLTNKQIGASAFRIQSDG 118

```



```

Db 1 MPIGNLGNVNGNHLIPAPLPSQTDGAAGGTGHLISSTGALGSRSLFSPLRNSMADS 60
QY 59 GNRASDVPLGVNPMRLAA--SEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQEDG 116
Db 61 VDSR--DIPGLPTNPSRLAAATSETCLLGGFEVLHDHKGPLDLINTQIGSPAFRVEQADG 118
QY 117 KHIAVGORNGVTSVVLSPQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEAR 176
Db 119 THAAIGKNGLEVSVTLSPQEWSSLSQIDTEGKRFVFTGGRGGGHPMVTVASDIAEAR 178
QY 177 QRILELLEPKGTG-----ESKGAGESKGVGELRESNAGENTTETQTSTSTSSLSRS 227
Db 179 TKILAKLDPDNHHGRQPKDVTRSVGVGSAGSI-----DDGV--VSETHTSTTSSVRS 230
QY 228 DKPLMALGTVAITGLIGLAATGIVQALALTPEDSPSTTTDPDAAASATETATRDQLTKEA 287
Db 231 DPKFWVSGAIAAGLAGLAGLAATGIAQALALTPEDPTTTDPDQANAASATKDLQTQEA 290
QY 288 FQNPQKVNIDELGNAPSGVLKDDVVAITEEQAAGEAKQQAIAENNAQAOKKYDEQ 347
Db 291 FKNPENQKVNIDANGNAIFSGELKDDIVQIAQAQAGEVARQQAQAVESNAQAQRDEQ 350
QY 348 QAKRQELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTSART 407
Db 351 HARRQELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTHT-----V 404
QY 408 VENKPAANTPAQGNVDTPGSDTMSRRSSMASTSTSTPTDTSIGTVQNPYADV-----K 462
Db 405 VQOQTGGNTPAQGGDTATRAEDASLNRDSQGSVASTHWSDS--SEVNPYAEVSGARNS 463
QY 463 TSLHD-----SQVPTSNSTSVQNMGTDSVVYSTIOHPPTDNDGARLLGNPSAGIQ 516
Db 464 LSAHQPEEHIYDEVAADPGYSVQNFSGGPVT-----GRLIGTPQGQIGQ 508
QY 517 STYARLALSGGLRHDHMGGLTGGSNSAVNTSNPPAPGSHRFV 558
Db 509 STYALLANSGLRLGMLGGLTGGESAVSSVNAATPGVRFV 550

RESULT 10
Q7BHL5
ID AC Q7BHL5 PRELIMINARY; PRT; 547 AA.
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Translocated intimin receptor Tir.
OS Citrobacter rodentium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=67825;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D8S100;
RX MEDLINE=21437640; PubMed=11553577;
RX DOI=10.1126/IAI.69.10.6323-6335.2001;
RA Deng W., Li Y., Vallance B.A., Finlay B.B.;
RT "Locus of enterocyte effacement from Citrobacter rodentium: sequence
RT analysis and evidence for horizontal transfer among attaching and
RT effacing pathogens."
RL Infect Immun. 69:6323-6335(2001).
DR EMBL; AF311901; AAL06376.1; -.
DR GO; GO:0004872; Fireceptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMNR.
KW Receptor.
SQ SEQUENCE 547 AA; 56283 MW; 02CAC6D625FA6EE1 CRC64;

```

Query Match 55.3%; Score 1569.5; DB 2; Length 547;

```

Best Local Similarity 57.0%; Pred. No. 4.5e-75;
Matches 332; Conservative 69; Mismatches 122; Indels 59; Gaps 11;

QY 1 MPIGNLGNVNNNSIPPAPLPSQTDGA--GGRQQLINSTGLGSRALFTFVRNSMADS 58
Db 1 MPIGNLGNNNISNLIIPAPLPSQTDGATRGNGSSLISSSTGLSRLFLPSLRSSIVDT 60
QY 59 GNRASDVPLGVNPMRLAASEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQEDGCH 118
Db 61 VDSR--DVPGLPEHPLRFATSETCLHGGFEVLHDHKGPLDLTLNKGIGSSVFRVEQPDGTH 118
QY 119 IAVGORNQGVTSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQ 178
Db 119 AAGVKDGVSVTLNSELQSLQSLDTEGNGRFVFTGGRGGSGHAMVTVASDISAREK 178
QY 179 ILELLEPKGTG-----ESKGAGESKGVGELRESNAGENTTETQTSTSTSSLSRSDP 229
Db 179 IIAKLDPDNHHGRQPKDITRSVGVGSAGMGD-----GV--VSETHTSTTSSVRSRDP 230
QY 230 KLWLALGTVAITGLIGLAATGIVQALALTPEDSPSTTTDPDAAASATETATRDQLTKEAQ 289
Db 231 KFWVSGAIAAGLAGLAGLAATGIVQALVALTPAPDDPTTTDPDDEAANAATAKDQLTKEAFQ 290
QY 290 NPDNOKVNIDELGNAPSGVLKDDVVAITEEQAAGEAKQQAIAENNAQAOKKYDEQQA 349
Db 291 NPDNOKVNIDELGNAPSGELKDDVVAIQADQAKVAGEARQQAQAVESNAQAQRHDDQQA 350
QY 350 KRQBELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTSARTVE 409
Db 351 KRQQLDLSSGIGYGLSSALIVGGGIGAGVTAALHRRNPPTTEQTATTHS-----VIQ 403
QY 410 NKPAANTPAQGNVDTPGSDTMSRRSSMASTST--FPDTSIGTVQNPYADV----- 460
Db 404 QQTGNTPAQGGADTGVENASLTRDSQASVASTQWSDTS--GDVNPYAEGWMSRNP 461
QY 461 ----VKTSLHDSQVPTSNSTSVQNMGTDSVVYSTIOHPPTDNDGARLLGNPSAGIQ 516
Db 462 SLLAQPEEPIYDEVADPN-----YSVQHFGSNPNVPTG--RLVGSFGQGIQ 505
QY 517 STYARLALSGGLRHDHMGGLTGGSNSAVNTSNPPAPGSHRFV 558
Db 506 STYALLANSGLRLGMLGGLTGGESAGSAANAATPGVRFV 547

RESULT 11
Q9WXK1
ID AC Q9WXK1 PRELIMINARY; PRT; 547 AA.
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MPEC;
RA Okutani A., Itoh K., Sasakawa C.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB026719; BAA77400.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMNR.
KW Receptor.
SQ SEQUENCE 547 AA; 56283 MW; 02CAC6D625FA6EE1 CRC64;

```



```
Query Match      55.3%; Score 1569.5; DB 2; Length 547;
Best Local Similarity 57.0%; Pred. No. 4.5e-75;
Matches 332; Conservative 69; Mismatches 122; Indels 59; Gaps 11;

QY 1 MPIGNLGNPNVNNIIPAPPLPSQTDGA--GGRQLINSTGPGSRALFTPVNRMSADS 58
DB 1 MPIGNLGNNNISNNLIIPAPPLPSQTDGATRGNGSSLISSGSLGSRLLFSLRSSIVDT 60

QY 59 GDNRASDPGLPVNPMRLAASEITLNDGFVLDHGGPLDTLNRQIGSSVFRVETQEDGKH 118
DB 61 VDNR--DVPGLEPHEPLRFATSETCLHGGFEVLDHGGPLDTLNRQIGSSVFRVETQEDGKH 118

QY 119 IAVGORNQVETSVLSDOEYARLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQR 178
DB 119 AAVGKDGVEVSVTLNSELQSLQSLDTGEGNRFVFTGGRGGAGHAMVTVASDITEARQR 178

QY 179 ILELLEPKGTG-----ESKGAGESKGVGELRESNGAENTTETQTSTSTSLRSDP 229
DB 179 IIAKLDPNHHGRQPKDIDTRSVGVSASGMGD-----GV--VSETHTSTTTSSVRSDP 230

QY 230 KLWALGTVATGLIGLAATGIQVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQ 289
DB 231 KFWVSVGAIAAGLAGLAATGIQVQALTPAPDPPTTDPDEAANAEEAATKDQLTKEAFQ 290

QY 290 NPDNQKVNIDELGNAIPSGVLKDDVVANIEQAAGBEAKQQAENNAQAQKKYDEOQA 349
DB 291 NPDNQKVNIDELGNAIPSGELKDDVVAQIADQAKVAGEQARQQAQVESNAQAQRHDDQA 350

QY 350 KRQBELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVE 409
DB 351 KRQBELDLSSGIGYGLSSALIVGGIGAGVTAMLHRRNPPTTEQTIATTHS-----VIQ 403

QY 410 NKPNANTPAQGNVDTPGSEDTMESRRSSMASTST--PFDTSIGTVQNPYAD----- 460
DB 404 QQTGNTRAQGGADTTGVENASLRRDSQASVASTQMSDTS--GDVVPYAEQMSNRNP 461

QY 461 ----VKTSLHDSQVPTSNSTSVQNMGNNTDSVVYSTIQHPPRDITDNGARLGNPSAGIQ 516
DB 462 SLLAPEPIYDEVAPDPN-----YSVQHFGSNPNPTVG-RLVGSFGQGIQ 505

QY 517 STYARLALSGGLRHDMGGLTGGNSAVNTSNPPAPGSHRFV 558
DB 506 STYALLASSGGLRLGMGGLTGGESAGSAANAATTPGVERFV 547

RESULT 12
Q9ETI1 PRELIMINARY; PRT; 547 AA.
AC Q9ETI1;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE Translocated intimin receptor Tir.
OS Citrobacter rodentium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=67825;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=CDC 1843-73T, and DBS100;
RX MEDLINE=20553330; PubMed=1101562;
RA Luperchio S.A., Newman J.V., Dangler C.A., Schrenzel M.D.,
RA Brenner D.J., Steigerwalt A.G., Schauer D.B.;
RT "Citrobacter rodentium, the causative agent of transmissible murine
RT colonic hyperplasia, exhibits clonality: synonymy of C. rodentium and
RL mouse-pathogenic Escherichia coli.";
RL J. Clin. Microbiol. 38:4343-4350(2000).
DR EMBL; AF301618; AAG40758.1; -
DR EMBL; AF301617; AAG25642.1; -
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
```

```
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor C; 1.
DR Pfam; PF03549; Tir_receptor M; 1.
DR Pfam; PF07490; Tir_receptor N; 1.
DR PRINTS; PR01370; TRNSINTIMNR.
KW Receptor.
SQ SEQUENCE 547 AA; 56270 MW; CB8318B301049C37 CRC64;

Query Match      55.2%; Score 1568.5; DB 2; Length 547;
Best Local Similarity 57.0%; Pred. No. 5.1e-75;
Matches 332; Conservative 69; Mismatches 122; Indels 59; Gaps 11;

QY 1 MPIGNLGNPNVNNIIPAPPLPSQTDGA--GGRQLINSTGPGSRALFTPVNRMSADS 58
DB 1 MPIGNLGNNNISNNLIIPAPPLPSQTDGATRGNGSSLISSGSLGSRLLFSLRSSIVDT 60

QY 59 GDNRASDPGLPVNPMRLAASEITLNDGFVLDHGGPLDTLNRQIGSSVFRVETQEDGKH 118
DB 61 VDNR--DVPGLEPHEPLRFATSETCLHGGFEVLDHGGPLDTLNRQIGSSVFRVETQEDGKH 118

QY 119 IAVGORNQVETSVLSDOEYARLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQR 178
DB 119 AAVGKDGVEVSVTLNSELQSLQSLDTGEGNRFVFTGGRGGAGHAMVTVASDITEARQR 178

QY 179 ILELLEPKGTG-----ESKGAGESKGVGELRESNGAENTTETQTSTSTSLRSDP 229
DB 179 IIAKLDPNHHGRQPKDIDTRSVGVSASGMGD-----GV--VSETHTSTTTSSVRSDP 230

QY 230 KLWALGTVATGLIGLAATGIQVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQ 289
DB 231 KFWVSVGAIAAGLAGLAATGIQVQALTPAPDPPTTDPDEAANAEEAATKDQLTKEAFQ 290

QY 290 NPDNQKVNIDELGNAIPSGVLKDDVVANIEQAAGBEAKQQAENNAQAQKKYDEOQA 349
DB 291 NPDNQKVNIDELGNAIPSGELKDDVVAQIADQAKVAGEQARQQAQVESNAQAQRHDDQA 350

QY 350 KRQBELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVE 409
DB 351 KRQBELDLSSGIGYGLSSALIVGGIGAGVTAMLHRRNPPTTEQTIATTHS-----VIQ 403

QY 410 NKPNANTPAQGNVDTPGSEDTMESRRSSMASTST--PFDTSIGTVQNPYAD----- 460
DB 404 QQTGNTRAQGGADTTGVENASLRRDSQASVASTQMSDTS--GDVVPYAEQMSNRNP 461

QY 461 ----VKTSLHDSQVPTSNSTSVQNMGNNTDSVVYSTIQHPPRDITDNGARLGNPSAGIQ 516
DB 462 SLLAPEPIYDEVAPDPN-----YSVQHFGSNPNPTVG-RLVGSFGQGIQ 505

QY 517 STYARLALSGGLRHDMGGLTGGNSAVNTSNPPAPGSHRFV 558
DB 506 STYALLASSGGLRLGMGGLTGGESAGSAANAATTPGVERFV 547

RESULT 13
ID O50190 PRELIMINARY; PRT; 549 AA.
AC O50190;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RX MEDLINE=98050926; PubMed=9390560; DOI=10.1016/S0092-8674(00)80437-7;
RA Kenny B., Devinney R., Stein M., Reinscheid D.J., Frey E.A.,
RA Finlay B.B.;
RT "Enteropathogenic E. coli (EPEC) transfers its receptor for intimate
```



```
RT adherence into mammalian cells.;"
RN Cell 91.511-520(1997).
RA [2]
RP SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RA Stein M.S., Kenny B., Finlay B.B.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF013122; AAB88410.1; -.
DR HSP; Q9KWH9; IFO2.
DR GO: 0004872; F:receptor activity; IEA.
DR GO: 0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR:receptor.
DR Pfam: PF07489; TIR:receptor_C; 1.
DR Pfam: PF03549; TIR:receptor_M; 1.
DR Pfam: PF07490; TIR:receptor_N; 1.
DR PRINTS; PRO1370; TRNSINTMINR.
KW Receptor.
SQ SEQUENCE 549 AA; 56843 MW; 40C8B8B234409A08 CRC64;

Query Match 53.9%; Score 1531.5; DB 2; Length 549;
Best Local Similarity 56.2%; Pred. No. 4.6e-73;
Matches 327; Conservative 61; Mismatches 137; Indels 57; Gaps 11;

QY 1 MPIGNLGNPNVNSIPPAPLPSTQDGA--GGRGOLINSTGPGSLRALTPTVRNSMADS 58
DB 1 MPIGNLGNVNGNHLIPAPLPSTQDGAARGTGHLSSTGALGSRSLFSPRNSMADS 60
QY 59 GDNRASDVPLGVNPMRLAA--SEITLNDGFVLDHDPGLDNLNRQIGSSVFRVETQEDG 116
DB 61 VDSR--DIPGLTPNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGSPAFRVEQADG 118
QY 117 KHIAVGQNGVTSVLSQVRLQSIIDPEKDKPVFTGGRGGAGHAWTVASDITEAR 176
DB 119 THAAIGEKNGLEVSITSPQEWSSLSQIDTEGKNRFVFTGGRGGHPMTVASDIAEAR 178
QY 177 QRILELLEPKGTG-----ESKAGESKGVGELRESNGAENTTETOTSTSSLSRS 227
DB 179 TKILALDNDNGGQPKDVTSTVGVGSAGI-----DDGV--VSEHTTSTSSVRS 230
QY 228 DPKLMALGTVATGLGLAATGIVQALALTPEPDSPTTTDPDAASATETATRDQLTKEA 287
DB 231 DPKFVWSVGAAGLAGLAGTGAQALALTPEPDDPTTTDPDAANAESATKDQLTQEA 290
QY 288 FQNPQKNVNDLGNALPSGLVDVVANIEEQAAGEAKQAQAIENNAQAQKYDQ 347
DB 291 FKNPENQKNVIDANGNAIPSGELKDDIVQIAQAEAGEVARQAQAVESNAQAQRYEDQ 350
QY 348 QAKROELKVSAGYGLSGALILGGGIGVAVTAALHKNQNPVEQTTTTTTTTSART 407
DB 351 HARRQELQLSGIGYGLSALIVAGGIGVAVTAALHKNQNPVEQTTTTTTTTSART 403
QY 408 VENKANNTPAQGNVDTPGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADV-----K 462
DB 404 VVQQTGGIPQHKVALMPQERRRFRDRDSQGSVASTHMSDSS--SEVNPYAEVGGARNS 462
QY 463 TSLHD-----SQVTSNNTSVQNMGTDSVVYSTIQHPDRTDNGARLGNPSAGIQ 516
DB 463 LSAHQPEEHIYDEVAADPGYSIQNFSGPVT-----GRLGTGQGGIQ 507
QY 517 STVARLALSGRLHDMGGLTGGSNAVNTSNPPAPGSHRFV 558
DB 508 STYALLANSGLRLGMLGGLTSGGETAVSSVNAAPTGPVRFV 549

RESULT 14
Q79C12 PRELIMINARY; PRT; 367 AA.
AC Q79C12;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to the hypothetical 38.4kDa protein of REPEC 84/110/1; unknown function (Fragment).

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC RDEC-1;
RA Krejany E.O.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U59503; AAB02941.1; -.
DR GO: 0004872; F:receptor activity; IEA.
DR GO: 0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR:receptor.
DR Pfam: PF07489; TIR:receptor_C; 1.
DR Pfam: PF03549; TIR:receptor_M; 1.
DR Pfam: PF07490; TIR:receptor_N; 1.
DR PRINTS; PRO1370; TRNSINTMINR.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 367 AA; 37797 MW; 00F84FCAC8FB5718 CRC64;

Query Match 40.4%; Score 1146.5; DB 2; Length 367;
Best Local Similarity 61.7%; Pred. No. 6.3e-53;
Matches 242; Conservative 38; Mismatches 83; Indels 29; Gaps 5;

QY 171 DITEARQRIILELLEPKGTGSGKAGESKGVGELRESNGAENTTETOTSTSSLSRDPK 230
DB 1 DIAEARQRIIDKLEPKDTKETKEPG-----DPNSGEGKIIIEHTSTSTSSLRDPK 51
QY 231 LMLALGTVATGLGLAATGIVQALALTPEPDSPTTTDPDAASATETATRDQLTKEAFQ 290
DB 52 LMLSGLTGAAGLIGNAAGIAQVALTPEPDPITTPDAANATAAQAADQLTKEAFQ 111
QY 291 PDNQKNVIDELGNALPSGLVDVVANIEEQAAGEAKQAQAIENNAQAQKYDQ 350
DB 112 PDNQKNVIDENGNALPSGELKDDVVQIAQAEQAAGEQAQAIENNAQAQKYDQ 171
QY 351 ROELKVSAGYGLSGALILGGGIGVAVTAALHKNQNPVEQTTTTTTTTSARTVEN 410
DB 172 REQEMSLSGVGYGSGALILGGGIGVAVTAALHKNQNPVEQTTTTTTTTSARTVEN 223
QY 411 KPANTPAQGNVDTPGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADVKTSLHDSQ 470
DB 224 QTNNASAGQNTDTSQSPESPASRRNSNASLASNSDTSSTGTVENPYADV-----GM 276
QY 471 PTSNNTSVQNMGTDSVV---YSTIQHPDRTDNGARLGNPSAGIQSTVARLALSG 526
DB 277 PRNDSLARISEPIYDEVAADPNYSVIOHFGSNPVTG-RLVGTGQGIQSTVALLASG 335
QY 527 GLRHDGMLTGGSNAVNTSNPPAPGSHRFV 558
DB 336 GLRLGMLTGGSNAVNTSNPPAPGSHRFV 367

RESULT 15
Q7YU77 PRELIMINARY; PRT; 1374 AA.
AC Q7YU77;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SD04153p.
GN Name=Alhambra;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champs M., Chavez C., Dorsett V., Drensek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
```

[illegible]

Search completed: May 13, 2005, 11:47:35  
Job time : 182 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:57:23 ; Search time 42 Seconds  
(without alignments)  
1278.308 Million cell updates/sec

Title: US-09-189-415b-11  
Perfect score: 558  
Sequence: 1 MPIGNLGNPNVNSIPPAP.....SNSAVNTSNPPAGSHRFV 558

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 8

Total number of hits satisfying chosen parameters: 210

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	558	100.0	558	2 A98199	translocated intim
2	558	100.0	558	2 E86045	probable transloca
3	13	2.3	166	2 C90029	hypothetical prote
4	12	2.2	139	2 D86417	probable auxin-ind
5	12	2.2	458	2 T31631	hypothetical prote
6	12	2.2	524	2 S33640	homeotic protein s
7	12	2.2	560	2 T32661	hypothetical prote
8	12	2.2	569	2 S47277	gp88 protein - mur
9	12	2.2	802	2 A36910	xylanase, beta(1,3
10	12	2.2	1023	2 S12519	glutactin - fruit
11	12	2.2	3712	2 S18253	laminin alpha-1 ch
12	12	2.2	4377	2 A55575	ankyrin 3, long sp
13	11	2.0	67	2 B56888	alkaline phosphata
14	11	2.0	108	2 T26880	hypothetical prote
15	11	2.0	164	2 T26561	hypothetical prote
16	11	2.0	183	2 S05358	hypothetical prote
17	11	2.0	208	2 T46896	merozoite surface
18	11	2.0	217	2 S01358	salivary glue prot
19	11	2.0	234	2 T25560	hypothetical prote
20	11	2.0	245	2 T26868	hypothetical prote
21	11	2.0	274	2 A45632	merozoite surface
22	11	2.0	278	2 S39310	merozoite surface
23	11	2.0	284	2 T22023	hypothetical prote
24	11	2.0	304	2 T15922	hypothetical prote
25	11	2.0	327	2 S20074	promastigote surfa
26	11	2.0	341	2 T32949	hypothetical prote
27	11	2.0	342	2 T29557	hypothetical prote
28	11	2.0	373	2 T29596	hypothetical prote
29	11	2.0	385	2 JC7783	RAD 23B protein -

30	11	2.0	415	2 T32467	hypothetical prote
31	11	2.0	468	2 A5476	protein kinase (EC
32	11	2.0	477	2 A54843	nemo, form 1 - fru
33	11	2.0	484	2 S58868	G protein-coupled
34	11	2.0	512	2 T02498	probable WRKY-type
35	11	2.0	516	2 S19252	1-aminocyclopropan
36	11	2.0	517	2 T20658	probable zinc meta
37	11	2.0	518	2 S31442	1-aminocyclopropan
38	11	2.0	519	2 T37339	hypothetical prote
39	11	2.0	525	2 A35596	nuclear pore glyco
40	11	2.0	526	2 A56573	nuclear pore compl
41	11	2.0	530	2 T32812	hypothetical prote
42	11	2.0	551	2 S18408	alkaline phosphata
43	11	2.0	559	2 B36307	alkaline phosphata
44	11	2.0	651	2 T21175	hypothetical prote
45	11	2.0	680	2 T19939	hypothetical prote
46	11	2.0	681	2 T23454	hypothetical prote
47	11	2.0	698	2 A54796	regulatory protein
48	11	2.0	732	2 T25937	hypothetical prote
49	11	2.0	770	2 T22808	hypothetical prote
50	11	2.0	781	2 S51592	hypothetical prote
51	11	2.0	825	2 T29634	hypothetical prote
52	11	2.0	831	2 T08611	hypothetical prote
53	11	2.0	889	2 A35679	rep protein - slim
54	11	2.0	975	2 T08606	protein phosphatas
55	11	2.0	1002	2 T30546	major surface glyco
56	11	2.0	1076	2 JC2217	major surface glyco
57	11	2.0	1083	2 JC2300	cell surface glyco
58	11	2.0	1089	2 T14576	nosa protein - eli
59	11	2.0	1093	2 T18275	1-phosphatidylinos
60	11	2.0	1099	2 T18257	phospholipase C -
61	11	2.0	1271	2 D64237	hypothetical prote
62	11	2.0	1282	2 JE0120	glycoprotein A - m
63	11	2.0	1402	2 T17456	cell surface prote
64	11	2.0	1635	2 T14075	chitinase (EC 3.2.
65	11	2.0	1671	2 S71628	sensory transducti
66	11	2.0	1737	2 A59235	unconventional myo
67	11	2.0	1832	2 T31113	mucin-like glycopr
68	11	2.0	1858	2 T18273	1-phosphatidylinos
69	10	1.8	127	2 T51538	nitrilase associat
70	10	1.8	187	2 T49491	hypothetical prote
71	10	1.8	195	2 T19617	hypothetical prote
72	10	1.8	263	2 S01360	salivary glue prot
73	10	1.8	307	1 GSFF3	salivary glue prot
74	10	1.8	388	2 T16861	hypothetical prote
75	10	1.8	390	2 T49619	hypothetical prote
76	10	1.8	393	2 B86189	protein T25N20.9 l
77	10	1.8	395	2 T45599	hypothetical prote
78	10	1.8	435	2 T25350	hypothetical prote
79	10	1.8	572	2 T16865	hypothetical prote
80	10	1.8	577	2 G89430	protein K02E2.3 (i
81	10	1.8	648	1 JQ1150	protein kinase (EC
82	10	1.8	649	2 T24505	hypothetical prote
83	10	1.8	662	2 A45155	mucin FIM-C.1 - Af
84	10	1.8	788	2 S05661	muscarinic acetyl
85	10	1.8	816	2 C69493	hypothetical prote
86	10	1.8	947	2 T08605	hypothetical prote
87	10	1.8	1008	2 T30544	major surface glyco
88	10	1.8	1014	2 T18759	hypothetical prote
89	10	1.8	1017	2 T30542	major surface glyco
90	10	1.8	1022	2 T30543	major surface glyco
91	10	1.8	1030	2 T18374	B-cell receptor pr
92	10	1.8	1390	2 T14004	trfa protein - eli
93	10	1.8	1513	2 T23681	hypothetical prote
94	10	1.8	1570	2 T18272	1-phosphatidylinos
95	9	1.6	124	2 T48833	hypothetical prote
96	9	1.6	202	2 F86755	prophage p12 prote
97	9	1.6	213	2 T23865	hypothetical prote
98	9	1.6	224	2 G86148	TiN6.14 protein -
99	9	1.6	232	2 A60095	larval glue protei
100	9	1.6	492	2 A41907	methyl-CpG-binding
101	9	1.6	500	1 EPFF	zip protein precur
102	9	1.6	622	2 T51223	hypothetical prote

103 1.6 645 2 T29818 hypothetical prote  
104 1.6 708 2 T29669 hypothetical prote  
105 1.6 906 2 S03313 hypothetical prote  
106 1.6 943 2 S6824 rngB protein, cyto  
107 1.6 1408 2 S16148 gene serrate prote  
108 1.6 1510 2 T33100 hypothetical prote  
109 1.6 3672 2 T23433 hypothetical prote  
110 1.6 3704 2 T37316 hypothetical laminin a  
111 1.6 76 2 B96809 protein F28K19.18  
112 1.4 78 2 A75442 hypothetical prote  
113 1.4 80 2 A91255 hypothetical prote  
114 1.4 80 2 E86095 hypothetical prote  
115 1.4 80 2 A65210 hypothetical 7.4 k  
116 1.4 81 2 AE1013 probable exported  
117 1.4 83 2 T20306 hypothetical prote  
118 1.4 105 2 T06085 probable auxin-ind  
119 1.4 147 2 T01039 hypothetical prote  
120 1.4 167 2 T33602 hypothetical prote  
121 1.4 169 2 T32698 hypothetical prote  
122 1.4 173 2 S16162 cruzipain (EC 3.4.  
123 1.4 176 2 T26212 hypothetical prote  
124 1.4 180 2 B45613 surface antigen FU  
125 1.4 193 2 T24370 hypothetical prote  
126 1.4 210 2 T49785 hypothetical prote  
127 1.4 225 2 S59852 DNA-binding protei  
128 1.4 240 2 AD3584 23S ribosomal RNA m  
129 1.4 270 2 T41759 LEF-1 orf14 - Bomb  
130 1.4 272 2 G71618 merozoite surface  
131 1.4 275 2 S09774 hypothetical prote  
132 1.4 286 2 B45632 merozoite surface  
133 1.4 287 2 S39615 merozoite 45K surf  
134 1.4 300 2 A39112 merozoite 45K surf  
135 1.4 302 2 A39615 merozoite 45K surf  
136 1.4 302 2 C86480 33.2K hypothetical  
137 1.4 303 2 T03571 probable peptidase  
138 1.4 314 2 D96703 hypothetical prote  
139 1.4 327 2 T49514 hypothetical prote  
140 1.4 333 2 T45720 probable transcrip  
141 1.4 333 2 F84775 hypothetical prote  
142 1.4 347 2 B39112 merozoite 45K surf  
143 1.4 381 2 T03640 protein phosphatas  
144 1.4 384 2 A44146 syndecan-3 - chick  
145 1.4 389 2 T33340 hypothetical prote  
146 1.4 390 2 T02845 DNA-damage inducib  
147 1.4 394 2 T20633 hypothetical prote  
148 1.4 406 2 T16520 hypothetical prote  
149 1.4 416 2 T14021 hypothetical prote  
150 1.4 422 2 T49513 gastric mucin rela  
151 1.4 434 2 T47772 hypothetical prote  
152 1.4 444 2 C87297 dihydroorotase [im  
153 1.4 467 2 A60667 cysteine proteinas  
154 1.4 467 2 A45629 cysteine proteinas  
155 1.4 475 2 T27811 conserved hypothet  
156 1.4 490 2 T02843 hypothetical prote  
157 1.4 496 2 F90181 nitrate reductase  
158 1.4 513 2 F83162 hal3 protein - yea  
159 1.4 531 2 S57752 probable exported  
160 1.4 540 2 AD0522 involved in polyke  
161 1.4 557 2 H69678 nuclear factor-kap  
162 1.4 558 2 JC2004 hypothetical prote  
163 1.4 587 2 D84426 hypothetical prote  
164 1.4 591 2 T25636 hydroxymethylgluta  
165 1.4 592 2 A32107 hypothetical prote  
166 1.4 592 2 T34446 hypothetical prote  
167 1.4 621 2 T20307 outer membrane pro  
168 1.4 634 2 A64521 transcription fact  
169 1.4 639 2 G02919 probable outer mem  
170 1.4 668 2 A71986 hypothetical prote  
171 1.4 672 2 T20310 hypothetical prote  
172 1.4 672 2 T32557 ABC transporter, b  
173 1.4 710 2 A99486 gelatinase B (EC 3  
174 1.4 712 2 I46031 hypothetical prote  
175 1.4 721 2 B90487

176 8 1.4 730 2 A53064 folded gastrulatio  
177 8 1.4 771 2 T21633 hypothetical prote  
178 8 1.4 786 1 A47547 serine proteinase  
179 8 1.4 790 2 T34293 hypothetical prote  
180 8 1.4 802 2 T21315 hypothetical prote  
181 8 1.4 816 2 S84439 hypothetical prote  
182 8 1.4 854 2 T03107 glycoprotein B - a  
183 8 1.4 892 2 T06818 DNA topoisomerase  
184 8 1.4 907 2 A45560 sporozoite surface  
185 8 1.4 918 2 D71407 hypothetical prote  
186 8 1.4 936 2 T06190 lipoxigenase (EC 1  
187 8 1.4 977 2 T16232 hypothetical prote  
188 8 1.4 1013 2 T33470 hypothetical prote  
189 8 1.4 1047 2 A55617 masquerade precurs  
190 8 1.4 1057 2 T04874 hypothetical prote  
191 8 1.4 1096 2 T08619 probable metallopr  
192 8 1.4 1129 2 T25635 hypothetical prote  
193 8 1.4 1172 2 T00065 hypothetical prote  
194 8 1.4 1175 2 T25634 hypothetical prote  
195 8 1.4 1229 2 T25697 hypothetical prote  
196 8 1.4 1269 2 T14476 p51A protein - eli  
197 8 1.4 1272 2 T30248 fragile X mental r  
198 8 1.4 1279 2 T13613 hypothetical prote  
199 8 1.4 1314 1 TNBXR6 transcription regu  
200 8 1.4 1335 2 T18289 racGAP protein - s  
201 8 1.4 1409 2 T29594 hypothetical prote  
202 8 1.4 1603 2 S17983 gene posterior sex  
203 8 1.4 1813 2 T30564 resistance protein  
204 8 1.4 1969 2 T08875 histidine kinase h  
205 8 1.4 3020 2 A43932 mucin 2 precursor,  
206 8 1.4 3133 2 S52093 hemocytin - silkwo  
207 8 1.4 3216 2 C90538 hypothetical prote  
208 8 1.4 3828 2 T13857 trithorax protein  
209 8 1.4 3848 2 T17414 tipC protein - eli  
210 8 1.4 5170 2 T15348 hypothetical prote

## ALIGNMENTS

## RESULT 1

A98199  
translocated intinin receptor Tir [imported] - Escherichia coli (strain O157:H7, substra  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: A98199  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: A98199  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-558 <HAY>  
A:Cross-references: UNIPROT:Q9R396; GB:BA000007; PIDN:BA037984.1; PID:gl31364036; GSPDB:G  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: EC04561

Query Match 100.0%; Score 558; DB 2; Length 558;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MPIGNLGNPNVNSIPPAPPLPSQTDGAGRGQLINSTGPLGSRALPTVRNSMADSGD 60  
Db 1 MPIGNLGNPNVNSIPPAPPLPSQTDGAGRGQLINSTGPLGSRALPTVRNSMADSGD 60  
Qy 61 NRASDVCLPNVWRLAASEITLNDGFEVLHDHGLDPLTNQIGSSVFRVETOEGKHIA 120  
Db 61 NRASDVCLPNVWRLAASEITLNDGFEVLHDHGLDPLTNQIGSSVFRVETOEGKHIA 120  
Qy 121 VQRNGVETSVLSQDEYARLQSIDPEGKOKFVTGGRGGAGHAMVTVASDITEARQIRL 180

```
Db 121 VQQRNGVTSVLSDOEYARLQSLIDPEGKDKFVFTGGGAGHAMVTVASDITEARQRL 180
QY 181 ELLEPKGTGESKGAGSKGVGLRESNSGAENTTETQTSTSTSSLRSDPKLMLALGTAT 240
Db 181 ELLEPKGTGESKGAGSKGVGLRESNSGAENTTETQTSTSTSSLRSDPKLMLALGTAT 240
QY 241 GLIGLAATGIVQALALTPPDSPTTDPDAAASATETATRDQLTKEAFQNPQKNWIDE 300
Db 241 GLIGLAATGIVQALALTPPDSPTTDPDAAASATETATRDQLTKEAFQNPQKNWIDE 300
QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIIENNAQAQKYDEQAKRQEELKVSSG 360
Db 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIIENNAQAQKYDEQAKRQEELKVSSG 360
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTTTTTSARTVENKPNNTPAQG 420
Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTTTTTSARTVENKPNNTPAQG 420
QY 421 NVDTPGSEDTMESRRSSMASTSTSTFDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
Db 421 NVDTPGSEDTMESRRSSMASTSTSTFDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
QY 481 NMGNNTDSVVYTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db 481 NMGNNTDSVVYTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
QY 541 SAVNTSNNPPAPGSHRFV 558
Db 541 SAVNTSNNPPAPGSHRFV 558

RESULT 2
E86045
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: E86045
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E86045
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <STO>
A:Cross-references: UNIPROT:O9R396; GB:AE005174; NID:G12518449; PIDN:AA658825.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
A:Gene: tir

Query Match 100.0%; Score 558; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGHNPVNNISIPAPPSPQTDGAGRGQLINSTGPGSRALFPVNSMADSGD 60
Db 1 MPIGNLGHNPVNNISIPAPPSPQTDGAGRGQLINSTGPGSRALFPVNSMADSGD 60
QY 61 NRASDVPGLPVNPMLAASEITLNDGFVLDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120
Db 61 NRASDVPGLPVNPMLAASEITLNDGFVLDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120
QY 121 VQQRNGVTSVLSDOEYARLQSLIDPEGKDKFVFTGGGAGHAMVTVASDITEARQRL 180
Db 121 VQQRNGVTSVLSDOEYARLQSLIDPEGKDKFVFTGGGAGHAMVTVASDITEARQRL 180
QY 181 ELLEPKGTGESKGAGSKGVGLRESNSGAENTTETQTSTSTSSLRSDPKLMLALGTAT 240
Db 181 ELLEPKGTGESKGAGSKGVGLRESNSGAENTTETQTSTSTSSLRSDPKLMLALGTAT 240
QY 241 GLIGLAATGIVQALALTPPDSPTTDPDAAASATETATRDQLTKEAFQNPQKNWIDE 300
```

```
Db 241 GLIGLAATGIVQALALTPPDSPTTDPDAAASATETATRDQLTKEAFQNPQKNWIDE 300
QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIIENNAQAQKYDEQAKRQEELKVSSG 360
Db 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIIENNAQAQKYDEQAKRQEELKVSSG 360
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTTTTTSARTVENKPNNTPAQG 420
Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTTTTTSARTVENKPNNTPAQG 420
QY 421 NVDTPGSEDTMESRRSSMASTSTSTFDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
Db 421 NVDTPGSEDTMESRRSSMASTSTSTFDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
QY 481 NMGNNTDSVVYTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db 481 NMGNNTDSVVYTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
QY 541 SAVNTSNNPPAPGSHRFV 558
Db 541 SAVNTSNNPPAPGSHRFV 558

RESULT 3
C90029
Hypothetical protein SA2097 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C90029
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A99758; MUID:21311952; PMID:11418146
A:Accession: C90029
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-166 <KUR>
A:Cross-references: UNIPROT:Q99RW9; GB:BA000018; PID:G13702104; PIDN:BA843396.1; GSPDB:G
A:Experimental source: strain N315
A:Genetics:
A:Gene: SA2097

Query Match 2.3%; Score 13; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 QTTTTTTTTTTTS 404
Db 43 QTTTTTTTTTTTS 55

RESULT 4
D86417
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86417
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Hestermann, R.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86417
A>Status: preliminary
```

A:Molecule type: DNA  
A:Residues: 1-139 <STO>  
A:Cross-references: UNIPROT:Q9C7Q5; GB:AE005172; NID:g10092232; PIDN:AAG12648.1; GSPDB:G  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T31631  
A:Map position: 1

Query Match 2.2%; Score 12; DB 2; Length 139;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404  
|||  
Db 22 TTTTITTTTTTTS 33

RESULT 5  
T31631  
hypothetical protein Y57A10A.i - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T31631  
R:Smyle, R.  
submitted to the EMBL Data Library, September 1999  
A:Reference number: Z21048  
A:Accession: T31631  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-458 <WIL>  
A:Cross-references: UNIPROT:Q9NA83; EMBL:AL117195; NID:el549729; PIDN:CAB55014.1; CESP:Y  
A:Experimental source: clone Y57A10A  
C:Genetics:  
A:Gene: CESP:Y57A10A.i  
A:Introns: 8/3; 54/3; 112/3; 151/1

Query Match 2.2%; Score 12; DB 2; Length 458;  
Best Local Similarity 100.0%; Pred. No. 0.0037;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404  
|||  
Db 135 TTTTITTTTTTTS 146

RESULT 6  
S33640  
homeotic protein smox-2, engrailed-like - fluke (Schistosoma mansoni)  
C:Species: Schistosoma mansoni  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 16-Aug-2004  
C:Accession: S33640; S27841  
R:Webster, P.J.; Mansour, T.E.  
Mech. Dev. 38, 25-32, 1992  
A:Title: Conserved classes of homeodomains in Schistosoma mansoni, an early bilateral me  
A:Reference number: S33640; MUID:92399260; PMID:1356008  
A:Accession: S33640  
A:Molecule type: mRNA  
A:Residues: 1-524 <WEB>  
A:Cross-references: UNIPROT:Q266601; EMBL:S44191; EMBL:M85305; NID:gl61103; PIDN:AAA29929  
C:Genetics:  
A:Gene: smox-2  
C:Superfamily: homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:424-480/Domain: homeobox homology <HOX>

Query Match 2.2%; Score 12; DB 2; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404  
|||  
Db 105 TTTTITTTTTTTS 116

RESULT 7

T32661  
hypothetical protein K11D12.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T32661  
R:Henkhaus, J.; Wohldmann, P.; Gillam, B.  
submitted to the EMBL Data Library, December 1997  
A:Description: The sequence of C. elegans cosmid K11D12.  
A:Reference number: Z21207  
A:Accession: T32661  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-560 <HEN>  
A:Cross-references: EMBL:AF039047; PIDN:AAB94223.1; GSPDB:GN00023; CESP:K11D12.1  
A:Experimental source: strain Bristol N2; clone K11D12  
C:Genetics:  
A:Gene: CESP:K11D12.1  
A:Map position: 5  
A:Introns: 5/3; 48/3; 90/3; 127/3; 149/3; 190/1; 207/1; 233/3; 264/1; 480/1

Query Match 2.2%; Score 12; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 0.0044;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404  
|||  
Db 348 TTTTITTTTTTTS 359

RESULT 8  
S47277  
gp88 protein - murine cytomegalovirus  
C:Species: murine cytomegalovirus, murine herpesvirus 1  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S47277  
R:Thaete, R.; Lucin, P.; Schneider, K.; Koszinowski, U.  
submitted to the EMBL Data Library, February 1994  
A:Reference number: S47277  
A:Accession: S47277  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-569 <THA>  
A:Cross-references: UNIPROT:Q83183; EMBL:X77798; NID:g535195; PIDN:CAAS4825.1; PID:g53519  
C:Superfamily: murine cytomegalovirus gp88 protein

Query Match 2.2%; Score 12; DB 2; Length 569;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 QTTTITTTTTTTT 403  
|||  
Db 472 QTTTITTTTTTTT 483

RESULT 9  
A36910  
xylanase, beta(1,3-1,4)-glucanase - Ruminococcus flavefaciens  
C:Species: Ruminococcus flavefaciens  
C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A36910  
R:Flint, H.J.; Martin, J.; McPherson, C.A.; Daniel, A.S.; Zhang, J.X.  
J. Bacteriol. 175, 2943-2951, 1993  
A:Title: A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-glucanase domain  
A:Reference number: A36910; MUID:93259938; PMID:8491715  
A:Accession: A36910  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-802 <FLI>  
A:Cross-references: UNIPROT:Q9S310; GB:S61204; NID:g385910; PIDN:AAB26620.1; PID:g385911  
A:Note: sequence extracted from NCBI backbone (NCBIN:131871, NCBI:P:131872)  
F:42-235/Domain: endo-1,4-beta-xylanase homology <XYL>  
F:259-401/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

```

Query Match          2.2%; Score 12; DB 2; Length 802;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTS 404
    |||||
Db 533 TTTTITTTTTTS 544

RESULT 10
glutactin - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S12519
R:Olson, P.F.; Fessler, L.I.; Nelson, R.E.; Sterne, R.E.; Campbell, A.G.; Fessler, J.H.
EMBO J. 9, 1219-1227, 1990
A:Title: Glutactin, a novel Drosophila basement membrane-related glycoprotein with sequ
A:Reference number: S12519; MUID:90214632; PMID:2108864
A:Accession: S12519
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1023 <OLS>
A:Cross-references: UNIPROT:P33438; EMBL:X53286; NID:G297084; PIDN:CAA37380.1; PID:G2970
C:Genetics:
A:Introns: 390/3

Query Match          2.2%; Score 12; DB 2; Length 1023;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTS 404
    |||||
Db 605 TTTTITTTTTTS 616

RESULT 11
laminin alpha-1 chain precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 16-Sep-1992 #sequence_revision 24-Jul-1997 #text_change 09-Jul-2004
R:Kusche-Gullberg, M.; Garrison, K.; MacKrell, A.J.; Fessler, J.H.
EMBO J. 11, 4519-4527, 1992
A:Title: Laminin A chain: expression during Drosophila development and genomic sequence.
A:Reference number: S28399; MUID:93049203; PMID:1425586
A:Accession: S28399
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-3712 <KUS>
A:Cross-references: UNIPROT:Q00174; GB:M96388; NID:g157799; PIDN:AAA28662.1; PID:g157800
R:Garrison, K.; MacKrell, A.J.; Fessler, J.H.
J. Biol. Chem. 266, 22899-22904, 1991
A:Title: Drosophila laminin A chain sequence, interspecies comparison, and domain struct
A:Reference number: S18253; MUID:92078147; PMID:1744083
A:Accession: S18253
A:Molecule type: mRNA
A:Residues: 1762-3712 <GAR>
A:Cross-references: EMBL:M75882; NID:g157797; PIDN:AAA28661.1; PID:g157798
C:Genetics:
A:Gene: FlyBase:LaNa
A:Cross-references: FlyBase:FBgn0002526
C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like H
C:Keywords: basement membrane; cell binding; coiled coil; disulfide bond; extracellular
F:273-330/Domain: laminin-type EGF-like homology <LEG>
F:333-400/Domain: laminin-type EGF-like homology <LE02>
F:541-584/Domain: laminin-type EGF-like homology <LEG1>
F:1776-2115/Domain: III <DOM3>
F:1776-1806/Domain: laminin-type EGF-like homology #status atypical <LE1>
F:1809-1856/Domain: laminin-type EGF-like homology <LE2>
F:1859-1914/Domain: laminin-type EGF-like homology <LE3>
F:1917-1967/Domain: laminin-type EGF-like homology <LE4>
F:1970-2014/Domain: laminin-type EGF-like homology <LE5>

```

```

F:2017-2061/Domain: laminin-type EGF-like homology <LE6>
F:2064-2109/Domain: laminin-type EGF-like homology <LE7>
F:2116-2697/Domain: I/II, heptad repeats <DOM2>
F:2698-3712/Domain: G <DOMG>
F:2698-2863/Domain: repeat G1 <RG1>
F:2864-3048/Domain: repeat G2 <RG2>
F:3049-3223/Domain: repeat G3 <RG3>
F:3079-3200/Domain: laminin G repeat homology <LG3>
F:3334-3528/Domain: repeat G4 <RG4>
F:3529-3712/Domain: repeat G5 <RG5>
F:1847,1850,1943,2024,2196,2215,2267,2301,2323,2482,2524,2538,2569,2699,2720,2890,2938,30

```

Query Match 2.2%; Score 12; DB 2; Length 3712;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTS 404  
 |||||  
Db 3273 TTTTITTTTTTS 3284

RESULT 12  
A:Accession: A55575  
N:Ankyrin 3, long splice form - human  
C:Species: Homo sapiens (man)  
C>Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 09-Jul-2004  
R:Kordeli, E.; Lambert, S.; Bennett, V.  
J. Biol. Chem. 270, 2352-2359, 1995  
A:Title: Ankyrin-G, A new ankyrin gene with neural-specific isoforms localized at the ax  
A:Reference number: A55575; MUID:95138209; PMID:7836469  
A:Accession: A55575  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-4377 <KOR>  
A:Cross-references: UNIPROT:Q12955; GB:U13616; NID:G608024; PIDN:AAA64834.1; PID:G608025  
C:Genetics:  
A:Gene: GDB:ANK3  
A:Cross-references: GDB:424503; OMIM:600465  
A:Map position: 10q21-10q21  
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology  
C:Keywords: alternative splicing; peripheral membrane protein  
F:73-105/Domain: ankyrin repeat homology <AN01>  
F:106-138/Domain: ankyrin repeat homology <AN02>  
F:139-171/Domain: ankyrin repeat homology <AN03>  
F:172-200/Domain: ankyrin repeat homology <AN04>  
F:201-233/Domain: ankyrin repeat homology <AN05>  
F:234-266/Domain: ankyrin repeat homology <AN06>  
F:267-299/Domain: ankyrin repeat homology <AN07>  
F:300-332/Domain: ankyrin repeat homology <AN08>  
F:333-365/Domain: ankyrin repeat homology <AN09>  
F:366-398/Domain: ankyrin repeat homology <AN10>  
F:399-431/Domain: ankyrin repeat homology <AN11>  
F:432-464/Domain: ankyrin repeat homology <AN12>  
F:465-497/Domain: ankyrin repeat homology <AN13>  
F:498-530/Domain: ankyrin repeat homology <AN14>  
F:531-563/Domain: ankyrin repeat homology <AN15>  
F:564-596/Domain: ankyrin repeat homology <AN16>  
F:597-629/Domain: ankyrin repeat homology <AN17>  
F:630-662/Domain: ankyrin repeat homology <AN18>  
F:663-695/Domain: ankyrin repeat homology <AN19>  
F:696-728/Domain: ankyrin repeat homology <AN20>  
F:729-761/Domain: ankyrin repeat homology <AN21>  
F:762-794/Domain: ankyrin repeat homology <AN22>  
F:795-827/Domain: ankyrin repeat homology <AN23>

Query Match 2.2%; Score 12; DB 2; Length 4377;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTS 404  
 |||||

```
Db      3968 TTTT TTTT TTTT TTTT TTTT 3979

RESULT 13
alkaline phosphatase (EC 3.1.3.1), intestinal type II - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 16-Aug-2004
C:Accession: B56888
R:Engle, M.J.; Alpers, D.H.
Clin. Chem. 38, 2506-2509, 1992
A:Title: The two mRNAs encoding rat intestinal alkaline phosphatase represent two distinct
A:Reference number: A56888; MUID:93092310; PMID:1458592
A:Accession: B56888
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-67 <ENG>
A:Experimental source: duodenal mucosa
A>Note: sequence extracted from NCBI backbone (NCBIN:121249, NCBIP:121252)
C:Superfamily: Alkaline phosphatase
C:Keywords: intestine; membrane protein; phosphoric monoester hydrolase

Query Match      2.0%; Score 11; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      393 TTTT TTTT TTTT TTTT TTTT 403
Db      27 TTTT TTTT TTTT TTTT TTTT 37

RESULT 14
hypothetical protein Y43F8C.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26880
R:Ainscough, R.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20279
A:Accession: T26880
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-108 <WIL>
A:Cross-references: UNIPROT:O9XWN0; EMBL:AL032637; PIDN:CAA21621.1; CESP:Y43F8C.9
A:Experimental source: clone Y43F8C
C:Genetics:
A:Gene: CESP:Y43F8C.9
A:Introns: 40/3

Query Match      2.0%; Score 11; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      393 TTTT TTTT TTTT TTTT TTTT 403
Db      50 TTTT TTTT TTTT TTTT TTTT 60

RESULT 15
hypothetical protein Y24F12A.d - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000
C:Accession: T26561
R:Rienhard, N.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20233
A:Accession: T26561
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-164 <WIL>
A:Cross-references: EMBL:AL110480; PIDN:CAB54380.1; CESP:Y24F12A.d
```

A:Experimental source: clone Y24F12A  
C:Genetics:  
A:Gene: CESP:Y24F12A.d  
A:Introns: 137/1  
C:Superfamily: Caenorhabditis elegans hypothetical protein Y9D1A.2

Query Match 2.0%; Score 11; DB 2; Length 164;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT TTTT 403  
Db 112 TTTT TTTT TTTT TTTT TTTT 122

Search completed: May 13, 2005, 12:07:18  
Job time : 48 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: May 13, 2005, 11:15:16 ; Search time 168 Seconds  
(without alignments)  
1284.597 Million cell updates/sec

Title: US-09-189-415b-11  
Perfect score: 2840  
Sequence: 1 MPIGNLGHNPVNNISIPAP.....SNSAVNTNPPAPGSHRFV 558

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 16Dec04: \*  
1: Geneseqp1980s: \*  
2: Geneseqp1990s: \*  
3: Geneseqp2000s: \*  
4: Geneseqp2001s: \*  
5: Geneseqp2002s: \*  
6: Geneseqp2003as: \*  
7: Geneseqp2003bs: \*  
8: Geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2840	100.0	558	ADC00799	Enterohae
2	2624.5	92.4	559	AAy06221	EHEC E. c
3	1534.5	54.0	549	AAy06220	EPEC E. c
4	406	14.3	107	AAy06220	EPEC E. c
5	187	6.6	1300	ABs20576	Intimin C
6	182.5	6.4	2586	ABb68075	Drosophil
7	180	6.3	596	ABb68078	Drosophil
8	180	6.3	596	AAy99408	Human PRO
9	180	6.3	596	AAb66157	Protein o
10	180	6.3	596	AAU29178	Human PRO
11	180	6.3	596	AAb87575	Human PRO
12	180	6.3	596	ABg95900	Human sec
13	180	6.3	596	ABu58554	Human PRO
14	180	6.3	596	ABu88102	Novel hum
15	180	6.3	596	ABu84417	Human sec
16	180	6.3	596	ABr66291	Human sec
17	180	6.3	596	ABr65681	Human sec
18	180	6.3	596	ABu99621	Human sec
19	180	6.3	596	ABu82860	Human PRO
20	180	6.3	596	ABu99981	Novel hum
21	180	6.3	596	ABr68230	Human sec
22	180	6.3	596	ABu96283	Novel hum
23	180	6.3	596	ABu92714	Human sec
24	180	6.3	596	ABu08791	Human sec
25	180	6.3	596	ABo02843	Human sec
				ABr74997	Human sec

26	180	6.3	596	6	ABR94759	Human PRO
27	180	6.3	596	6	ABU85732	Human PRO
28	180	6.3	596	6	ABU98892	Novel hum
29	180	6.3	596	6	ABU98107	Novel hum
30	180	6.3	596	6	ABU91813	Novel hum
31	180	6.3	596	6	ABU89506	Human PRO
32	180	6.3	596	6	ABU86347	Human sec
33	180	6.3	596	6	ABU67560	Human sec
34	180	6.3	596	6	ABU80588	Human PRO
35	180	6.3	596	6	ABU90925	Novel hum
36	180	6.3	596	6	ABO33984	Human sec
37	180	6.3	596	6	ABR99506	Human sec
38	180	6.3	596	6	ABR98896	Human sec
39	180	6.3	596	6	ABO16419	Human sec
40	180	6.3	596	6	ABR92319	Human sec
41	180	6.3	596	6	ABO18960	Human sec
42	180	6.3	596	6	ABR78381	Human sec
43	180	6.3	596	6	ABU72001	Novel hum
44	180	6.3	596	6	ABU85117	Novel hum
45	180	6.3	596	6	ABO00256	Novel hum

ALIGNMENTS

RESULT 1

ADC00799  
ID ADC00799 standard; protein; 558 AA.  
XX  
AC ADC00799;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Enterohaemorrhagic E. coli O157:H7-specific protein SEQ ID NO: 844.  
XX  
KW enterohaemorrhagic; anti-bacterial.  
XX  
OS Escherichia coli; O157:H7.  
XX  
PN JP2002355074-A.  
XX  
PD 10-DEC-2002.  
XX  
PF 24-JAN-2002; 2002JP-00015959.  
XX  
PR 24-JAN-2001; 2001JP-00112010.  
XX  
PA (UYTS-) UNIV TSUKUBA.  
XX

DR WPI; 2003-451640/43.  
XX  
PT Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule and a polypeptide and its use, a polypeptide, a vector and a host cell.  
XX  
PS Claim 3; SEQ ID NO 844; 2067pp; Japanese.

XX  
The invention relates to a novel enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule. A polynucleotide of the invention has anti-bacterial activity. The polypeptide can be used in detection and/or treatment of O157:H7 infection. The nucleotide sequence of the genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present invention represents an E. coli O157:H7-specific polypeptide of the

XX  
Sequence 558 AA;

Query Match 100.0%; Score 2840; DB 7; Length 558;  
Best Local Similarity 100.0%; Pred. No. 1.4e-193;  
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGHNPVNNISIPAPLPSTQDAGGGRGQLINSTGLGSRALFTPVNRNSMADSGD 60  
DB 1 MPIGNLGHNPVNNISIPAPLPSTQDAGGGRGQLINSTGLGSRALFTPVNRNSMADSGD 60

QY 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120  
 DB 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120  
 QY 121 VQORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQIL 180  
 DB 121 VQORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQIL 180  
 QY 181 ELLEPKGTGESKGAGESKGVGELRESNCAENTTQTSTSSLSRSDPKLWALGTAT 240  
 DB 181 ELLEPKGTGESKGAGESKGVGELRESNCAENTTQTSTSSLSRSDPKLWALGTAT 240  
 QY 241 GLIGLAATGIVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQNPQKNVIDE 300  
 DB 241 GLIGLAATGIVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQNPQKNVIDE 300  
 QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQAKRQEELKVSSG 360  
 DB 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQAKRQEELKVSSG 360  
 QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQVPEOTTTTTTTTTTSARTVENKPNNTPAQG 420  
 DB 361 AGYGLSGALILGGGIGVAVTAALHRKNQVPEOTTTTTTTTTTSARTVENKPNNTPAQG 420  
 QY 421 NVDTPGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480  
 DB 421 NVDTPGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480  
 QY 481 NMGNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDHMGGLTGGSN 540  
 DB 481 NMGNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDHMGGLTGGSN 540  
 QY 541 SAVNTSNPPAPGSHRFV 558  
 DB 541 SAVNTSNPPAPGSHRFV 558

RESULT 2

ID AAY06221 standard; protein; 559 AA.  
 XX AAY06221;  
 AC AAY06221;  
 DT 16-AUG-1999 (first entry)  
 DE EHEC E. coli translocated intimin receptor (Tir).  
 XX Tir; translocated intimin receptor; Hp90; enterohaemorrhagic; EHEC;  
 KW infection; diagnosis; vaccine.  
 XX Escherichia coli.

Key Location/Qualifiers  
 Misc-difference 453

FT /note= "encoded by codon of 1 apparent nucleotide,  
 FT causing frameshift in the DNA sequence"

PN WO9224576-A1.

XX 20-MAY-1999.

PF 10-NOV-1998; 98WO-CA001042.

PR 12-NOV-1997; 97US-0065130P.

XX (UYBR-) UNIV BRITISH COLUMBIA.

XX Finlay BB, Kenny B, Devinney R, Stein M;

XX WPI: 1999-337712/28.

DR N-PSDB; AAX58859.

XX

PT New translocated intimin receptor useful for treating infection by  
 PT enteropathogenic or enterohaemorrhagic Escherichia coli.  
 XX Claim 7; Page 55-58; 91pp; English.  
 PS  
 CC The present sequence represents Tir, a novel translocated intimin  
 CC receptor (formerly termed Hp90) from an enterohaemorrhagic Escherichia  
 CC coli (EHEC) strain. The sequence was deduced from an isolated Tir  
 CC polynucleotide (see AAX58859). Tir proteins are secreted by attaching and  
 CC effacing pathogens such as EHEC and EPEC (see AAY06220) E. coli. The  
 CC bacterial pathogens insert their own receptors into mammalian cell  
 CC surfaces, to which the pathogen then adheres to trigger additional host  
 CC signaling events and actin nucleation. Diagnosis of disease caused by  
 CC pathogenic E. coli can be performed by use of antibodies that bind to Tir  
 CC to detect the protein or the use of nucleic acid probes for detection of  
 CC nucleic acids encoding Tir polypeptide. Isolated Tir nucleic acids, Tir  
 CC peptides, a recombinant method for producing recombinant Tir, antibodies  
 CC which bind to Tir, and a kit for the detection of Tir-producing E. coli  
 CC are provided. A method of immunising a host with Tir to induce a  
 CC protective immune response is also provided. In addition, Tir fusion  
 CC proteins can be used in attenuated E. coli to induce a cell-mediated  
 CC immune response to other polypeptides, e.g. antigens. A method for  
 CC screening for compounds which interfere with the binding of bacterial  
 CC pathogens to their receptors is further provided  
 XX  
 SQ Sequence 559 AA;

Query Match 92.4%; Score 2624.5; DB 2; Length 559;

Best Local Similarity 94.0%; Pred. No. 3.1e-170;

Matches 528; Conservative 3; Mismatches 24; Indels 7; Gaps 3;

QY 1 MPIGNLGNPNVNNISIPAPPPLPSQTDGAGRGQLNSTGPLGSRALTPPVNSMADSGD 60

DB 1 MPIGNLGNPNVNNISIPAPPPLPSQTDGAGRGQLNSTGPLGSRALTPPVNSMADSGD 60

QY 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120

DB 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120

QY 121 VQORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQIL 180

DB 121 VQORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQIL 180

QY 181 ELLEPKGTGESKGAGESKGVGELRESNCAENTTQTSTSSLSRSDPKLWALGTAT 240

DB 181 ELLEPKGTGESKGAGESKGVGELRESNCAENTTQTSTSSLSRSDPKLWALGTAT 240

QY 241 GLIGLAATGIVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQNPQKNVIDE 300

DB 241 GLIGLAATGIVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQNPQKNVIDE 300

QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQAKRQEELKVSSG 360

DB 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQAKRQEELKVSSG 360

QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQVPEOTTTTTTTTTTSARTVENKPNNTPAQG 420

DB 361 AGYGLSGALILGGGIGVAVTAALHRKNQVPEOTTTTTTTTTTSARTVENKPNNTPAQG 420

QY 421 NVDTPGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSL-HDSQVPTSNSNTSV 479

DB 421 NVDTPGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSL-HDSQVPTSNSNTSV 479

QY 480 QNM---GNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDHMGGLT 536

DB 478 RLFRIWGIQISVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDHMGGLT 537

QY 537 GGSNSAVNTSNPPAPGSHRFV 558

DB 538 GGSNSAVNTSNPPAPGSHRFV 559

RESULT 3



PS Claim 8; Page 76; 96pp; English.  
 CC The present invention describes a method of screening for an inhibitor of  
 CC intinin binding to eukaryotic cells. The method comprises exposing an  
 CC intinin polypeptide having a Tir-independent cell binding activity to  
 CC test agents, and obtaining an inhibitor based on its ability to bind the  
 CC polypeptide. The inhibitors are used in the prevention, treatment and/or  
 CC diagnosis of bacterial infections, preferably by enteropathic and/or  
 CC enterohaemorrhagic Escherichia coli, Shiga toxinogenic E. coli, Hafnia  
 CC alvei or Citrobacter freundii, or especially E. coli O157:H7. The  
 CC infections cause a histopathological effect known as attachment and  
 CC effacement on intestinal epithelial cells. The inhibitors can be used to  
 CC produce food supplements or additives, especially where the food is a  
 CC milk substitute. The method can be used to sort cells based on their  
 CC ability to bind to a Tir independent cell binding domain of an intinin  
 CC polypeptide. Polypeptides having Tir-independent intinin binding activity  
 CC can be used to produce a vaccine against a bacterial disease. The present  
 CC sequence represents a specifically claimed intinin C-terminal Tir binding  
 CC domain amino acid sequence, for use in the method of the present  
 CC invention  
 XX  
 SQ Sequence 107 AA;

Query Match 14.3%; Score 406; DB 3; Length 107;  
 Best Local Similarity 72.0%; Pred. No. 3.6e-21;  
 Matches 77; Conservative 18; Mismatches 12; Indels 0; Gaps 0;  
 QY 252 QALATTPEDSPSTTTDPDAASATATATRDQLTKEAFQNPDKVNIIDELGNALIPSGYLK 311  
 DB 1 QALATTPEDDPTTDPQANAASATKQDTQAFKNPENQKVNIDANGNALIPSGELK 60  
 QY 312 DDVANIIEQAKAEEAKQAIENNAQAQKYDEQQAQKEELKVS 358  
 DB 61 DDIVQIAQAKEGEVARQAVESNAQAQRYEDQHARRQEELQLS 107

RESULT 5  
 ABB68075  
 ID ABB68075 standard; protein; 1300 AA.  
 XX  
 AC ABB68075;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 31017.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 PN WO200171042-A2.  
 PD 27-SEP-2001.  
 PF 23-MAR-2001; 2001WO-US009231.  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABLI2178.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 31017; 21pp + Sequence Listing; English.  
 XX

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA  
 CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1300 AA;

Query Match 6.6%; Score 187; DB 4; Length 1300;  
 Best Local Similarity 20.4%; Pred. No. 0.00039;  
 Matches 130; Conservative 74; Mismatches 245; Indels 188; Gaps 25;  
 QY 5 NLGNPNVNSIPAPPLPSPQTDG---AGRGQLINSTGLSGSRALFTPVNSMADSGDN 61  
 DB 438 NSTSNSNTNDSTGPSETSTNGLVASGGAG-----GATGAAMLPTP---SQOSTGK 488  
 QY 62 RAS-----DVPGLPVNPMRLA-----ASEITLNDGFVLHDHGLDPLTLNRQ 102  
 DB 489 EATAAVALLEKLVNVVSPITMKELRQKMGTKYDAEMINAAAYQQQHH----- 539  
 QY 103 IGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEARLQSIDPEGDKKFVFTGGRGAG 162  
 DB 540 --QHHFHHHHHHHHHGHQASTGAETAATAVQMAAMQKPG-----VGTGAAG 588  
 QY 163 HAMTVASDITEARQRIELLELPKGTGSKGAGSKGVGELRESNGAENTTQTSTST 222  
 DB 589 NAGATTVSSVA-----AGAGSEVNGRSTSLRKSMRVNS-----TSSSI 627  
 QY 223 SSLRSDPKMLALGTATGLGLAATGIVQALALTPEPDSPTTTDPDAASATATATRDQ 282  
 DB 628 STASADEVI-----APVVAASISLPSKAPVVLMPRCCKPAQMAIALHQ 670  
 QY 283 LTKEAFQNPDKVNIIDELGNALIPSGVLKDDVANIIEQAKAEEAKQK-----A 333  
 DB 671 SQQRQLRSEHQEKLTID-----GESSD---TSSEQKK---EQQDQDLPQKMF 716  
 QY 334 IENNAQAQKYDEQQAQKEELKVSAGYGLSGALILGGGIGVAVTAALHRKQVPEQT 393  
 DB 717 LAEPFQPEKS--EEKQEQKQKRVTRNSAGRVGL-----VARLATAINN-----IA 760  
 QY 394 TTTTITTTTTSARTV---ENKPAANTPAQGNVDPGSEDWTMESRRSSMASTSTFFDTSS 450  
 DB 761 TTTTSSSSSNKATTTTCNNHNSNNSRINNSLRLSVKSRKPAPSEASSIPSTSS 820  
 QY 451 IGTVQ-----NPFYADVKTSLHDSQVPTSNSTSVQNMGNNTDSVVYSTIOH----- 495  
 DB 821 ENQQQATRRSCSPTPAYKKNLLASFPDPPSTQGIKEQLKDESVTYSPVKQKSRRAA 880  
 QY 496 -----PPRDTTNGARL-LGNPSAGIQST-----YARLALSGG 527  
 DB 881 LAAQSIHCEALGGFTGSGQRKRAQAGEPTTSCSTTISNVNVEPLKTPERRKLTLR 940  
 QY 528 LRH-----DMGLTGGNSAVNTSNPPAPGSHR 556  
 DB 941 MKRSPILDEVIELG--TSLNNGGARG-----APGSHR 971

RESULT 6  
 ABB66878  
 ID ABB66878 standard; protein; 2586 AA.  
 XX  
 AC ABB66878;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 27426.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;



```

PR 23-SEP-1998; 98US-0103474P.
PR 23-SEP-1998; 98US-0104175P.
PR 23-SEP-1998; 98US-0104176P.
PR 23-SEP-1998; 98US-0104177P.
PR 23-SEP-1998; 98US-0104178P.
PR 23-SEP-1998; 98US-0104179P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101741P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101915P.
PR 24-SEP-1998; 98US-0101916P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102307P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102484P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 07-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 08-OCT-1998; 98US-0103711P.
PR 14-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106022P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 29-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-NOV-1998; 98US-0106464P.
PR 30-NOV-1998; 98US-0108856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 03-NOV-1998; 98US-0106934P.
PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.

PR 17-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
XX (GETH ) GENENTECH INC.
XX
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX WPI; 2000-237871/20.
XX N-PSDB; AAA37090.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or secreted
XX PRO polypeptides, useful for screening of potential peptide or small
XX molecule inhibitors of the relevant receptor/ligand interactions.
XX
XX Claim 12; Fig 138; 773pp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
XX receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
XX transmembrane and receptor PRO proteins can be used for screening of
XX potential peptide or small molecule inhibitors of the relevant
XX receptor/ligand interactions. The polypeptides and nucleotide sequences
XX encoding then have various industrial applications, including uses as
XX pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR
XX primers and hybridisation probes used in the isolation of the PRO
XX polypeptides from the present invention
XX
XX Sequence 596 AA;

Query Match 6.3%; Score 180; DB 3; Length 596;
Beat Local Similarity 19.2%; Pred. No. 0.00044;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNIGHNPVNSIIPPAPLPSTGAGGRGQLINSTPLGSRALFTPVNRNSMADSGDNR 63
DB 13 GLLHLEATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNMRLAASEITLNDGFEVLHDHGLDPLNRQIGSSV-FRVETOEDGRIHVG 122
DB 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFTTSSGISTATNSESTASSGSIATN 107
QY 123 QRNGVETS--VVLSDQEVARIQSIDPEGKQFVTGGRGGAGHAMVTVASDITEARQIL 180
DB 108 SESSTTSSGASTATNSE----SSTPSSGASTVTNNGSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGESKGAGESKGVCELRESNAGENTTETOTSTSTSLRSDPKLWALGTAT 240
DB 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDSPITTPDPAASATATATRDQLTKEAFQNPQNKVNDI 300
DB 212 N----SESTVSSRASTATNSESTT---SSGASTATNSESTTSSGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAGEEAKQAIENNAQAQKYDEQQAQR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QEELKVSSGAGYGL-SGALLIGGGIGVAVTAALHRKNQPVQOTTTTTTTTTTARTVEN 410
DB 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTVSSGISTVTNSESTSSGANTATN 362
QY 411 KPANNTPAQGNVDPGSEDNTMESRRSSMASTSTFFDTSSICTQNPYADVKT----- 464
DB 363 SESSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGAGTATNSESTTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGNNTDSVVYSTIQHPDRTDNGARLLCNPSAGQSTVARIAL 524
DB 422 NSDSTTSSSEASTATNSESTVSSGISTVTNSESTTSSGANTATNSESTTSSGSSVTS----- 474

```

QY 525 SGGLRDMGCLTGGNSAVNTNPPAPG 553  
 Db 475 AGSGTAALTGHTTSHA-STAVSEAKPG 502

RESULT 8  
 AAB66157  
 ID AAB66157 standard; protein; 596 AA.  
 XX  
 AC AAB66157;  
 XX  
 DT 02-APR-2001 (first entry)  
 DE Protein of the invention #69.  
 XX  
 KW Secreted; transmembrane; gene therapy.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200078961-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 18-FEB-2000; 2000WO-US004342.  
 XX  
 PR 23-JUN-1999; 99US-0141037P.  
 PR 20-JUL-1999; 99US-0144758P.  
 PR 26-JUL-1999; 99US-0145698P.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 29-OCT-1999; 99US-0162506P.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000376.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;  
 PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
 PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;  
 PI Williams PM, Wood WI;  
 XX  
 DR WPI; 2001-071395/08.  
 XX  
 PT Secreted and transmembrane proteins and nucleic acids designated PRO,  
 PT useful as hybridization probes, in chromosome and gene mapping and gene  
 PT therapy.  
 XX  
 PS Claim 1; Fig 138; 787pp; English.  
 XX  
 CC The present invention relates to secreted and transmembrane proteins.  
 CC These proteins and the DNA encoding them may be used as hybridization  
 CC probes, in chromosome and gene mapping and in the generation of anti-  
 CC sense RNA and DNA. They may also be used to generate either  
 CC transgenic animals or knockout animals which are in turn useful for  
 CC development and screening of therapeutically useful reagents. The nucleic  
 CC acids may also be used in gene therapy  
 XX  
 SQ Sequence 596 AA;

Query Match 6.3%; Score 180; DB 4; Length 596;  
 Best Local Similarity 19.2%; Pred. No. 0.00044;  
 Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNIGHNPNVNSIPPAPLPQTDGAGRGQLINSTGPLGSRALFFPVNSWADSGNRA 63  
 Db 13 GLLLHLEAATNS-----NETSANTGSSVISSG-----ASTATNSG 49  
 QY 64 SDYPGLVPVPMRLAAEITINDGFEVLHDHGLDITLNRQIGSSV-FRVEIQEDGKHIAVG 122  
 Db 50 SSVTSSGVSATISGSSVTSN-GVSIV-TNSEPHITSSGISTATNSEFTASSGISIATN 107

QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKDKVFVFTGRCGAGHAMVTVASDITEARQRL 180  
 Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVTNNGSSVTSSTSSLRSDPKLMLALGTVA 163  
 QY 181 ELLEPKGTGESKGAGESKGVGELRESNCAENTTETOTSTSTSSLRSDPKLMLALGTVA 240  
 Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTSSG-----ASTAT 211  
 QY 241 GLIGLAATGIVQALALTPEDSPFTTDPDAASATETATRDQITKEAFQNPQKNIDE 300  
 Db 212 N-----SESSTVSSRASTATNSESSTT---SSGASTATNSESRTTNGAGTATNSESSTTS 264  
 QY 301 LG-----NAIPSGVLKDDVAVNIEBOAKAAGEAKQAENNAQAKKYDEQOAKR 351  
 Db 265 SGASTATNDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302  
 QY 352 QEELKVSSGAGYGL--SGALLLGGGIGVAVTAALHRKNQPVQETTTTTTTTTTTSARTVEN 410  
 Db 303 SESSTTSSGASTATNDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSGANTATN 362  
 QY 411 KPANNTPAQGNVDTEGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADVKTSS----- 464  
 Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTAT 421  
 QY 465 LHDSQVPTNSNTSVQNMGNDSVYVSIQHPPTDNGARLLGNPSAGIOSTYARLAL 524  
 Db 422 NSDSSTTSSGASTATNSESSTVSSGISTVTNSESSTTSSGANTATNNGSSVTS----- 474  
 QY 525 SGGLRDMGCLTGGNSAVNTNPPAPG 553  
 Db 475 AGSGTAALTGHTTSHA-STAVSEAKPG 502

RESULT 9  
 AAU29178  
 ID AAU29178 standard; protein; 596 AA.  
 XX  
 AC AAU29178;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Human PRO polypeptide sequence #155.  
 XX  
 KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;  
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200168848-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006520.  
 XX  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 03-MAR-2000; 2000US-0187202P.  
 PR 06-MAR-2000; 2000US-0186968P.  
 PR 14-MAR-2000; 2000US-0189320P.  
 PR 14-MAR-2000; 2000US-0189328P.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 21-MAR-2000; 2000US-0190828P.  
 PR 21-MAR-2000; 2000US-0191007P.  
 PR 21-MAR-2000; 2000US-0191048P.  
 PR 21-MAR-2000; 2000US-0191314P.  
 PR 28-MAR-2000; 2000US-0192655P.  
 PR 29-MAR-2000; 2000US-0193032P.  
 PR 29-MAR-2000; 2000US-0193053P.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 04-APR-2000; 2000US-0194449P.











```
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 18-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0016897B.

Query Match
Beat Local Similarity 6.3%; Score 180; DB 6; Length 596;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHPNVNNSIPAPPLPSCDTCAGGCGQLINSTGLGSLAFTPVNRNSMAUDSGDNR 63
DB 13 GULLHLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPLGPNMRLAASEITLNDGFVLHDHGLDNLNRQIGSSV-FRVETQEDGKHIAVG 122
DB 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNEFFHTTSGISTATNSEFSTASSGISIATN 107
QY 123 QRNGVETS--VVLSDQEVARIQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQIL 180
DB 108 SESSTTSSGASTATNSE----SSTPSSGASTVNSGSSVTSSGASTATNSSESSTVSSRAS 163
QY 181 ELLEPKGTGESKAGESKVGELRESNCAENTTETQSTSTSSLSRSDPKLWLALGTVAT 240
DB 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEPDSPTTTPDPAASAATETATRDQLTKEAFQNPDKVNI 300
DB 212 N-----SESTVSSRATSTATNSESTT---SSGASTATNSESTTSSNGAGATATNSSESSTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEEAKQAIENNAQAQKXYDEQQAQR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST---ATN 302
QY 352 QEELKVSSGAGVGL-SGALLGGGIGVAVTAALHRKNPQVQTTTTTTTTTTSARTVEN 410
DB 303 SESSTTSSGASTATNSDSTSSGAGTATNSESTTVSSGISTVNTNSESTPSSGANTATN 362
QY 411 KPAANTPAQGNVDTPGSEDTMESRRSSMASSTSTPFDTSSTGTVQNPYADVKTSS-----464
DB 363 SESSTTSSGANTAT-NSSESTVSSGASTATNSESTTSSGVSTATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNGTSVONMGNTDSVVYSTIOHPPRDTTNGARLLGNPAGIQSTYARLAL 524
DB 422 NSDSSTTSSSEASTATNSSESSTVSSGISTVNTNSESTTSSGANTATNSGSSVTS-----474
QY 525 SGGLRHDHMGGLTGGNSAVNTSNPNPAG 553
DB 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502
```

RESULT 13  
ABU88102  
ID ABU88102 standard; protein; 596 AA.  
XX  
AC ABU88102;  
XX  
DT 07-JUL-2003 (first entry)  
XX  
DE Novel human secreted and transmembrane protein PRO1342.  
XX  
KW Human; secreted and transmembrane protein: PRO; gene therapy;  
KW tumour necrosis factor-alpha release; TNF-alpha release;  
KW chondrocyte proliferation; chondrocyte differentiation; tumour;  
KW adrenal tumour; lung tumour; colon tumour; breast tumour;  
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.  
XX  
OS Homo sapiens.  
XX  
PN US2003032127-A1.  
XX  
PD 13-FEB-2003.  
XX  
PF 26-JUN-2002; 2002US-00183012.  
XX  
PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0082250P.  
PR 21-OCT-1997; 97US-0083486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0083544P.  
PR 28-OCT-1997; 97US-0083564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0086120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 15-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 21-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.

```
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.

PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096953P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 15-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.

Query Match 6.3%; Score 180; DB 6; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.00044;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGNPNVNNISPPAPLPSTODCAGGRGQLINSTGPIGSRALFTPVNNSMADSGDRA 63
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLAAASEITLNDGFVLHDHGLDPLTNLROIGSSV-FRVETQEDGKHIAG 122
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHITSSGISTATNSEFSTASSGISIATN 107
QY 123 QNNGVETS--VVLSDQEVARLOSIDPEGKDKVFFTGGRGAGHAMVTVASDITEARQRL 180
```

Db 108 SSSSTTSGASTATNSE-----SSTPSSGASTVTNCGSSVTSSGASTATNSESSTVSSRAS 163  
 QY 181 ELLEPKGTGCKGAGESKGVGLRBSNGAENTTETOTSTSSLRSDPKLWALGTVAT 240  
 Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTSSG-----ASTAT 211  
 QY 241 GUGUAATGIVQALALTPEPDSPTTTPDPAASAATETATRDQLTKEAFQNPQNKVNIDE 300  
 Db 212 N-----SESSTVSSRASTATNSESSTT---SSGASTATNSESRTTNGAGTATNSESSTTS 264  
 QY 301 LG-----NAIPSGVLKDDVVANIEBQAKAAGEAKQQAENNAQAQKYDEQQAKR 351  
 Db 265 SGASTATNDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302  
 QY 352 OELKVSQAGYGL-SCALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTTTSARTVEN 410  
 Db 303 SESSTTSSGASTATNDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATN 362  
 QY 411 KPANNTPAQGNVDTPGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKT----- 464  
 Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTAT 421  
 QY 465 LHDSQVPTSNSTVQNMGNTDSVVYTIQHPPRTTNGARLLGNPAGIQSTVARLAL 524  
 Db 422 NSDSSTTSEASTATNSESSTVSSGISTVTNSESSTTSSGANTATNCGSVTS----- 474  
 QY 525 SGLRHDGMGLTGGNSAVNTSNPPAPG 553  
 Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 14

ID ABU84417 standard; protein; 596 AA.  
 XX AC ABU84417;  
 XX DT 02-AUG-2003 (first entry)  
 XX DE Human secreted/transmembrane protein (PRO) #155.  
 KW Human; secreted and transmembrane protein; PRO; TNF-alpha;  
 KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;  
 KW tissue typing.  
 XX OS Homo sapiens.  
 XX PN US2003032112-A1.  
 XX PD 13-FEB-2003.  
 XX PF 21-JUN-2002; 2002US-00176756.  
 XX PR 18-SEP-1997; 97US-0059263P.  
 PR 18-SEP-1997; 97US-0059266P.  
 PR 17-OCT-1997; 97US-0062250P.  
 PR 21-OCT-1997; 97US-0063486P.  
 PR 24-OCT-1997; 97US-0063120P.  
 PR 24-OCT-1997; 97US-0063121P.  
 PR 28-OCT-1997; 97US-0063540P.  
 PR 28-OCT-1997; 97US-0063541P.  
 PR 28-OCT-1997; 97US-0063544P.  
 PR 28-OCT-1997; 97US-0063564P.  
 PR 29-OCT-1997; 97US-0063734P.  
 PR 31-OCT-1997; 97US-0063870P.  
 PR 31-OCT-1997; 97US-0064103P.  
 PR 13-NOV-1997; 97US-0065311P.  
 PR 21-NOV-1997; 97US-0066120P.  
 PR 24-NOV-1997; 97US-0066466P.  
 PR 24-NOV-1997; 97US-0066772P.  
 PR 11-DEC-1997; 97US-0069335P.  
 PR 12-DEC-1997; 97US-0069425P.  
 PR 17-DEC-1997; 97US-0069870P.

PR 18-DEC-1997; 97US-0068017P.  
 PR 10-MAR-1998; 98US-0077450P.  
 PR 11-MAR-1998; 98US-0077632P.  
 PR 11-MAR-1998; 98US-0077649P.  
 PR 20-MAR-1998; 98US-0078886P.  
 PR 20-MAR-1998; 98US-0078939P.  
 PR 27-MAR-1998; 98US-0079664P.  
 PR 27-MAR-1998; 98US-0079786P.  
 PR 31-MAR-1998; 98US-0080107P.  
 PR 31-MAR-1998; 98US-0080194P.  
 PR 01-APR-1998; 98US-0080327P.  
 PR 01-APR-1998; 98US-0080333P.  
 PR 08-APR-1998; 98US-0081049P.  
 PR 08-APR-1998; 98US-0081070P.  
 PR 09-APR-1998; 98US-0081195P.  
 PR 15-APR-1998; 98US-0081838P.  
 PR 21-APR-1998; 98US-0082568P.  
 PR 21-APR-1998; 98US-0082569P.  
 PR 22-APR-1998; 98US-0082704P.  
 PR 22-APR-1998; 98US-0082797P.  
 PR 28-APR-1998; 98US-0083322P.  
 PR 29-APR-1998; 98US-0083495P.  
 PR 29-APR-1998; 98US-0083496P.  
 PR 29-APR-1998; 98US-0083499P.  
 PR 29-APR-1998; 98US-0083559P.  
 PR 05-MAY-1998; 98US-0084366P.  
 PR 06-MAY-1998; 98US-0084414P.  
 PR 07-MAY-1998; 98US-0084639P.  
 PR 07-MAY-1998; 98US-0084640P.  
 PR 07-MAY-1998; 98US-0084643P.  
 PR 15-MAY-1998; 98US-0085579P.  
 PR 15-MAY-1998; 98US-0085580P.  
 PR 15-MAY-1998; 98US-0085582P.  
 PR 15-MAY-1998; 98US-0085700P.  
 PR 18-MAY-1998; 98US-0086023P.  
 PR 22-MAY-1998; 98US-0086392P.  
 PR 22-MAY-1998; 98US-0086486P.  
 PR 28-MAY-1998; 98US-0087098P.  
 PR 28-MAY-1998; 98US-0087208P.  
 PR 02-JUN-1998; 98US-0087609P.  
 PR 02-JUN-1998; 98US-0087759P.  
 PR 03-JUN-1998; 98US-0087827P.  
 PR 04-JUN-1998; 98US-0088025P.  
 PR 04-JUN-1998; 98US-0088028P.  
 PR 04-JUN-1998; 98US-0088029P.  
 PR 04-JUN-1998; 98US-0088033P.  
 PR 04-JUN-1998; 98US-0088326P.  
 PR 05-JUN-1998; 98US-0088167P.  
 PR 05-JUN-1998; 98US-0088202P.  
 PR 05-JUN-1998; 98US-0088212P.  
 PR 05-JUN-1998; 98US-0088217P.  
 PR 09-JUN-1998; 98US-0088655P.  
 PR 10-JUN-1998; 98US-0088722P.  
 PR 10-JUN-1998; 98US-0088738P.  
 PR 10-JUN-1998; 98US-0088740P.  
 PR 10-JUN-1998; 98US-0088811P.  
 PR 10-JUN-1998; 98US-0088824P.  
 PR 10-JUN-1998; 98US-0088825P.  
 PR 10-JUN-1998; 98US-0088826P.  
 PR 11-JUN-1998; 98US-0088861P.  
 PR 11-JUN-1998; 98US-0088863P.  
 PR 12-JUN-1998; 98US-0088876P.  
 PR 12-JUN-1998; 98US-0089090P.  
 PR 12-JUN-1998; 98US-0089105P.  
 PR 16-JUN-1998; 98US-0089512P.  
 PR 16-JUN-1998; 98US-0089514P.  
 PR 17-JUN-1998; 98US-0089538P.  
 PR 17-JUN-1998; 98US-0089598P.  
 PR 17-JUN-1998; 98US-0089633P.  
 PR 18-JUN-1998; 98US-0089908P.  
 PR 19-JUN-1998; 98US-0089952P.  
 PR 22-JUN-1998; 98US-0090246P.  
 PR 22-JUN-1998; 98US-0090252P.

PR 22-JUN-1998; 98US-0090254P.  
 PR 24-JUN-1998; 98US-0090422P.  
 PR 24-JUN-1998; 98US-0090435P.  
 PR 24-JUN-1998; 98US-0090444P.  
 PR 24-JUN-1998; 98US-0090461P.  
 PR 24-JUN-1998; 98US-0090535P.  
 PR 24-JUN-1998; 98US-0090540P.  
 PR 25-JUN-1998; 98US-0090678P.  
 PR 25-JUN-1998; 98US-0090678P.  
 PR 25-JUN-1998; 98US-0090688P.  
 PR 25-JUN-1998; 98US-0090690P.  
 PR 25-JUN-1998; 98US-0090694P.  
 PR 25-JUN-1998; 98US-0090695P.  
 PR 25-JUN-1998; 98US-0090698P.  
 PR 26-JUN-1998; 98US-00105413.  
 PR 26-JUN-1998; 98US-0090862P.  
 PR 26-JUN-1998; 98US-0090863P.  
 PR 26-JUN-1998; 98US-0091010P.  
 PR 01-JUL-1998; 98US-0091359P.  
 PR 01-JUL-1998; 98US-0091544P.  
 PR 02-JUL-1998; 98US-0091478P.  
 PR 02-JUL-1998; 98US-0091486P.  
 PR 02-JUL-1998; 98US-0091626P.  
 PR 02-JUL-1998; 98US-0091632P.  
 PR 24-JUL-1998; 98US-0094008P.  
 PR 04-AUG-1998; 98US-0095282P.  
 PR 10-AUG-1998; 98US-0095998P.  
 PR 10-AUG-1998; 98US-0096012P.  
 PR 17-AUG-1998; 98US-0096757P.  
 PR 17-AUG-1998; 98US-0096766P.  
 PR 17-AUG-1998; 98US-0096867P.  
 PR 17-AUG-1998; 98US-0096891P.  
 PR 17-AUG-1998; 98US-0096897P.  
 PR 18-AUG-1998; 98US-0096949P.  
 PR 18-AUG-1998; 98US-0096959P.  
 PR 18-AUG-1998; 98US-0097022P.  
 PR 26-AUG-1998; 98US-0097952P.  
 PR 26-AUG-1998; 98US-0097954P.  
 PR 26-AUG-1998; 98US-0097955P.  
 PR 26-AUG-1998; 98US-0097971P.  
 PR 26-AUG-1998; 98US-0097974P.  
 PR 26-AUG-1998; 98US-0098014P.  
 PR 01-SEP-1998; 98US-0098716P.  
 PR 01-SEP-1998; 98US-0098723P.  
 PR 02-SEP-1998; 98US-0098803P.  
 PR 02-SEP-1998; 98US-0098821P.  
 PR 02-SEP-1998; 98US-0098843P.  
 PR 09-SEP-1998; 98US-0098602P.  
 PR 10-SEP-1998; 98US-0099741P.  
 PR 10-SEP-1998; 98US-0099754P.  
 PR 10-SEP-1998; 98US-0099763P.  
 PR 10-SEP-1998; 98US-0099812P.  
 PR 15-SEP-1998; 98US-0100388P.  
 PR 16-SEP-1998; 98US-0100662P.  
 PR 16-SEP-1998; 98US-0100664P.  
 PR 16-SEP-1998; 98US-0101751P.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 17-SEP-1998; 98US-0100683P.  
 PR 17-SEP-1998; 98US-0100684P.  
 PR 17-SEP-1998; 98US-0100919P.  
 PR 17-SEP-1998; 98US-0100930P.  
 PR 18-SEP-1998; 98US-0100849P.  
 PR 18-SEP-1998; 98US-0101014P.  
 PR 18-SEP-1998; 98US-0101068P.  
 PR 23-SEP-1998; 98US-0101471P.  
 PR 23-SEP-1998; 98US-0101472P.  
 PR 23-SEP-1998; 98US-0101475P.  
 PR 23-SEP-1998; 98US-0101477P.  
 PR 24-SEP-1998; 98US-0101738P.  
 PR 24-SEP-1998; 98US-0101739P.  
 PR 24-SEP-1998; 98US-0101743P.  
 PR 24-SEP-1998; 98US-0101922P.

PR 25-SEP-1998; 98US-0101786P.  
 PR 29-SEP-1998; 98US-0102207P.  
 PR 29-SEP-1998; 98US-0102240P.  
 PR 29-SEP-1998; 98US-0102330P.  
 PR 29-SEP-1998; 98US-0102331P.  
 PR 30-SEP-1998; 98US-0102487P.  
 PR 30-SEP-1998; 98US-0102570P.  
 PR 30-SEP-1998; 98US-0102571P.  
 PR 01-OCT-1998; 98US-0102684P.  
 PR 01-OCT-1998; 98US-0102687P.  
 PR 02-OCT-1998; 98US-0102965P.  
 PR 06-OCT-1998; 98US-0103258P.  
 PR 06-OCT-1998; 98US-0103449P.  
 PR 07-OCT-1998; 98US-00168978.  
 Query Match 6.3%; Score 180; DB 6; Length 596;  
 Best Local Similarity 19.2%; Pred. NO. 0.0004;  
 Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;  
 QY 4 GNIGHNPNVNSIPPAPLPQTDGAGRGQLINSTGLSRALETPVNSMADSGDNRA 63  
 Db 13 GLLHLLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49  
 QY 64 SDVPGLPVNPMRLAASEITLNDGFEVLHDGHPLOTNLRQIGSSV-FRVEQTQDGHIAVG 122  
 Db 50 SSVTSSGVSTATISGSSVTSN-GVSIIV-TNSEFHTTSGISTATNSESTASSGISIATN 107  
 QY 123 QRNGVETS--VVLSDQYARLQSIDPECKDFVFTGGRGGAGHAMVTVASDITEARQIL 180  
 Db 108 SESSTTSSGASTATNSE---SSTPSSGASTVTVNSGSSVTSSGASTATNSESTVSSRAS 163  
 QY 181 ELLEPKGTGSGKAGESKVGCELRESNGAENTTETQTSTSSLSRSPKLWLALGTVAT 240  
 Db 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211  
 QY 241 GLIGLAATGIQALALTPEDPSPTTDPDAAASATETATRDQLTKEAFQNPDKVKNIDE 300  
 Db 212 N----SESSTVSSRASTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264  
 QY 301 LG-----NAIPSGVLKDDVVANIEEOAKAAGEEAKQQAENNAQAQKYDQQAQR 351  
 Db 265 SGASTATNSDSSSTVSSGA---STATNSESTTSSGAST-----ATN 302  
 QY 352 QBELKVSSGAGYGL-SGALILGGGIGVAVTAAHLRKHPQVEQTITTTTTTTTSARTVEN 410  
 Db 303 SESSTTSSGASTATNSDSTTSSGAGTATNSESTVSSGISTVTVNSSESTSSGANTATN 362  
 QY 411 KPAANTPAQGNVDTPGSEDTWESRSSMASTSTTFDTTSGTGVQNPYADVKTST----- 464  
 Db 363 SESSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGASTATNSESTTSSGASTAT 421  
 QY 465 LHDSQVPTNSNTSYQNMGNSTDVVYVSTIOHPRDTTONGARLLGNPSAGIQSTYARLAL 524  
 Db 422 NSDSTTSSSEASTATNSESTVSSGISTVTVNSSESTTSSGANTATNSGSSVTS----- 474  
 QY 525 SGGLRHDMMGLTGGNSAVNTSNPPAPG 553  
 Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 15

ABR66291  
 ID ABR66291 standard; protein; 596 AA.

XX ABR66291;

XX 05-AUG-2003 (first entry)

XX Human secreted polypeptide PRO1342, SEQ ID NO:310.

XX Human; PRO; secreted protein; transmembrane protein;  
 KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;  
 KW chondrocyte; proliferation; differentiation; cartilage disorder;





```

PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 02-SEP-1998; 98US-0098960P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match      6.3%; Score 180; DB 6; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.00044;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLCNPNVNSIPPAPLPQTDGAGRGQLINSTGLGSRALFTPVNSMADSGDNEA 63
DB 13 GULLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49

QY 64 SDVPGPLVNPMLAASEITLNDGFEVLHDHGLDNLNROIGSSV-PRVETQEDGKHIAVG 122
DB 50 SSVTSSGVATISGSSVTSN-GVSIV-TNSEPHITSSGISTATNSEFSTASSGSIATN 107

QY 123 QRNGVETS--VVLSDQFYARLQSIDPEGKDKFVFTGCGGAGHAMVTVASDITEARQIL 180
DB 108 SSSSTTSSGASTATNSE---SSTPSSGASTVTSNGSSVTSSTSGASTATNSSTVSSRAS 163

QY 181 ELLEPKGTGSGKAGSKGVGELRESNGAENTTETOTSTSTSLRSDPKLWALGTAT 240
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSSESSTSSG-----ASTAT 211

QY 241 GLIGLAATGIVQALATTPEDSPPTTDPDAAASATETATRDOLTKEAFQNPONKWNIDE 300
DB 212 N-----SESSTVSRASATATNSESSTT---SSGASTATNSESRTTNGAGTATNSESSTTS 264

QY 301 LG-----NAIPSGVLKDDVIANIEBQAKAAGEAKQQAIIENNAQAQKYDEQOAKR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302

QY 352 QBELKVSSGAGYGL-SCALILGGIGVAVTAALHRKNQPVQTTTTTTTTTTTTTSARTVEN 410

```

```

DB 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESSTVSSGISTVTSNESSTPPSGANTATN 362
QY 411 KPANNTPAQGNVDTPGSDTMESSRRSSMASTSTFFDTSSTCTVQNPYADVKTSS----- 464
DB 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNTDSVVYSTIOHPPTTONGARLLGNPSAGIOSTYARLAL 524
DB 422 NSDSSTTSEASTATNSESSTVSSGISTVTSNESSTTSSGASTATNNGSSVTS----- 474
QY 525 SGGLRHDMMGGLTGGNSAVNTSNPPAPG 553
DB 475 AGSGTAALTGMHTTSHA-STAVSEAKPG 502

```

Search completed: May 13, 2005, 11:44:31  
Job time : 172 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**